ANALYSIS OF KALMAN FILTERING UNDER UNCERTAINTY IN NOISE COVARIANCES

By

SUWANCHAI SANGSUK-IAM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1987

TO MY PARENTS AND ALL OF MY TEACHERS

ACKNOWLEDGEMENTS

Professor T.E. Bullock, the chairman of my supervisory committee, has been a constant source of encouragement during my entire graduate study. His arrangement of financial support made my graduate study at the University of Florida possible. To him, I would like to express my deep gratitude.

I would like to express my sincere appreciation to the other members of my supervisory committee, Professors G. Basile, H.A. Latchman, Z.R. Pop-Stojanovic, and M.C.K. Yang. I would also like to acknowledge the Air Force Armament Laboratory, Eglin Air Force Base, Florida, for its financial support in terms of research assistantships during my graduate study.

A doctorate is much more than a dissertation. Earning one is a gradual educational process. I gratefully acknowledge Professors Kothom Ariya, Mongkol Dejnakarintra, Vichai Sungkachantranont, Suriyan Tishyadhigama of Chulalongkorn University, Bangkok, Thailand, Professor P.P. Khargonekar of the University of Minnesota, Professor E.W. Kamen of the University of Pittsburgh, Professors D.G. Childers, W.C. Parr, P.Z. Peebles, Jr., A.J. Rosalsky, J.R. Smith, and A. Vince of the University of Florida, and Professors D.G. Messerschmitt, S.S. Sastry, J. Walrand, and F.F. Wu of the University of California at Berkeley for their guidance and help. Finally, I wish to express my special gratitude to my parents for their love and understanding.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ANALYSIS OF KALMAN FILTERING
UNDER UNCERTAINTY IN NOISE COVARIANCES

By

Suwanchai Sangsuk-Iam

December, 1987

Chairman: Dr. T.E. Bullock
Major Department: Electrical Engineering

The Kalman filter has been widely accepted to be useful in many areas of applications. To implement the Kalman filter for a system of interest, covariances of noise disturbing the system must be completely specified. In most practical cases, the noise covariances are either inexactly known, unknown, or subject to abrupt changes. In this dissertation, Kalman filtering under the above uncertainties in noise covariances is analyzed. The dissertation consists of three main parts. The first part considers situations in which the Kalman filter is designed by using incorrect noise covariances. Behavior of the Kalman filter under incorrect noise covariances is then analyzed in this part. The second part discusses direct estimation of the noise covariances when they are unknown. Finally, cases in which the noise covariances are subject to abrupt changes are investigated in the third part.

CHAPTER I
INTRODUCTION

In 1960, Kalman published his celebrated paper on a new approach to linear filtering and prediction problems (Kalman, 1960). In his paper, Kalman formulated and solved the classical Wiener filtering problem from the state-space point of view. The solution was given in terms of recursive equations which are known today as the discrete-time Kalman filter. The continuous-time version of the filter was later given by Kalman and Bucy (1961). Since then, the Kalman filter has been extensively studied and widely used in many areas of applications. To implement the Kalman filter for a system of interest, both statistical and dynamical model parameters of the system must be completely specified. The exact values of the model parameters, however, are hardly known in most circumstances. Consequently, considerable research has been done on the effect of errors in model parameters.

Heffes (1966), Nishimura (1966, 1967a, 1967b, 1970), and Aasnaes and Kailath (1974) studied the effect of errors in statistical model parameters, while the effect of errors in dynamical model parameters was studied by Neal (1967), Koussoulas and Leondes (1986). Fagin (1964), Huddle and Wismer (1968), Griffin and Sage (1968, 1969), Fujita and Fukao (1970), Lainiotis and Sims (1970), and Brown and Sage (1971a, 1971b) investigated the effect of errors in both statistical and dynamical model parameters. A class of error phenomena known as divergence was examined by Schlee et al. (1967), Price (1968), and

Fitzgerald (1971). Bounds on the filter performance in the presence of modeling errors were derived in Toda and Patel (1978, 1980) and Patel and Toda (1984). In spite of considerable research mentioned above, few useful qualitative results have been obtained. In particular, no qualitative results were given at all in Fagin (1964), Heffes (1966), Neal (1967), Griffin and Sage (1968, 1969), Lainiotis and Sims (1970), and Koussoulas and Leondes (1986). They merely developed equations for evaluating the actual mean-squared-error performance of the Kalman filter in terms of parameter errors.

Inaccuracies in the initial state error covariance and noise covariances are quite common. The initial state error covariance, for example, is often poorly known. The selection of the initial state error covariance is usually based on physical intuition and common sense. However, under certain conditions (Jazwinski, 1970; Aasnaes and Kailath, 1974) the effect of errors in the initial state error covariance is lessened as more and more measurements are processed. The noise covariances, on the other hand, are either inexactly known, unknown, or subject to abrupt changes in most practical cases.

In this dissertation, Kalman filtering under the above uncertainties in noise covariances is analyzed. Particular emphasis is given to discrete-time Kalman filtering. The dissertation comprises three main parts. In the first part, we consider situations in which the Kalman filter is designed by using incorrect noise covariances. Behavior of the Kalman filter under incorrect noise covariances is then analyzed. The results developed in the first part provide useful insights in behavior of the Kalman filter when inexact values of noise covariances are used. Situations in which the noise covariances are

unknown are considered in the second part. A direct technique for estimating the unknown noise covariances, which is referred to as the stationary preprocessed measurement correlation (SPMC) technique, is discussed and analyzed in this part. In the third part, we consider cases in which the noise covariances are subject to abrupt changes. In particular, we are interested in abrupt increases in measurement noise covariances which can be used to model the effect of sensor failures in the form of increased inaccuracies. A scheme for detecting abrupt increases in measurement noise covariances is developed.

The organization of the dissertation is as follows. Chapter II contains analysis of discrete-time Kalman filtering under incorrect noise covariances. Similar analysis for continuous-time Kalman filtering is given in Chapter III. Discussion and analysis of direct estimation of noise covariances when they are unknown are given in Chapter IV. In Chapter V, cases in which noise covariances are subject to abrupt changes are considered, and a scheme for detecting abrupt increases in noise covariances is formulated. Conclusions of results presented in the dissertation are given in Chapter VI. The dissertation also contains four appendices which provide the relevant support material.

CHAPTER II
BEHAVIOR OF THE DISCRETE-TIME KALMAN FILTER UNDER INCORRECT NOISE COVARIANCES

In most practical cases, noise covariances are not completely known. Thus, the values of the noise covariances used in designing the Kalman filter are usually the approximated values of the actual noise covariances. It has been shown by Fitzgerald (1971) that incorrect values of the noise covariances can cause the filter to diverge. Because of common errors in noise covariances and possible filter divergence, it is important to investigate and understand behavior of the Kalman filter under incorrect noise covariances.

In this chapter, we analyze behavior of the discrete-time Kalman filter under incorrect noise covariances. In particular, we are interested in the characteristic of the actual performance of the Kalman filter. The quantity used to represent the filter performance is the actual one-step predictor error covariance. For simplicity of presentation, the initial state error covariance is assumed to be exactly known throughout the chapter. However, it should be noted that only minor and rather straightforward modifications of the results presented here are required in case that the initial state error covariance is incorrect as well.

The investigation given here is restricted to linear time-invariant systems with stationary noise processes. But for completeness of presentation, relevant results for linear time-varying systems with

nonstationary noise processes will be mentioned. Portions of the results contained in this chapter have been reported in Sangsuk-Iam and Bullock (1987a).

2.1 Preliminaries

Consider the discrete stochastic dynamical system described by

$$x_{k+1} = F_k x_k + w_k \tag{2.1}$$

$$y_k = H_k x_k + v_k \tag{2.2}$$

where x_k and y_k denote the state and the measurement, respectively. w_k and v_k denote respectively the system noise and the measurement noise, assumed white and zero-mean. Furthermore, x_o, $\{w_k\}$, and $\{v_k\}$ are assumed to be mutually uncorrelated.

Given $\{y_o, y_1, \ldots, y_{k-1}\}$, the linear minimum variance estimate of x_k can be computed recursively by the standard Kalman filter. Let us now consider the case that the filter has been designed with system and measurement noise covariances $Q_k \geq 0$ and $R_k > 0$, respectively, but yet the actual noise covariances are, respectively, Q_k^o and R_k^o. It should be noted that the estimate $\hat{x}_{k|k-1}$ of x_k given $\{y_o, y_1, \ldots, y_{k-1}\}$, computed from the filter, is no longer the linear minimum variance estimate. However, $\hat{x}_{k|k-1}$ is still an unbiased estimate of x_k .

Let M_k denote the one-step predictor error covariance computed from the filter. That is M_k satisfies the following Riccati difference equation (RDE):

$$M_{k+1} = F_k M_k F_k' - F_k M_k H_k' (H_k M_k H_k' + R_k)^{-1} H_k M_k F_k' + Q_k \quad ; M_o = \Pi \tag{2.3}$$

where the prime symbol denotes matrix transposition, and Π denotes the covariance of x_o .

Or equivalently,

$$M_{k+1} = \bar{F}_k M_k \bar{F}_k' + K_k R_k K_k' + Q_k \tag{2.4}$$

where $K_k := F_k M_k H_k' (H_k M_k H_k' + R_k)^{-1}$ and $\bar{F}_k := F_k - K_k H_k$.

Let us define the actual one-step predictor error covariance M_k^o as the error covariance associated with $\hat{x}_{k|k-1}$, i.e.,

$$M_k^o := \text{COV}(\tilde{x}_{k|k-1}, \tilde{x}_{k|k-1}) \quad \text{where} \quad \tilde{x}_{k|k-1} := x_k - \hat{x}_{k|k-1} .$$

It is straightforward to show that (Heffes, 1966)

$$M_{k+1}^o = \bar{F}_k M_k^o \bar{F}_k' + K_k R_k^o K_k' + Q_k^o \quad ; M_o^o = \Pi \tag{2.5}$$

The following simple but useful result, which follows immediately from (2.4) and (2.5), is due to Nishimura (1970).

<u>Theorem 2.1</u>: If $Q_k \geq (\leq)\ Q_k^o$ and $R_k \geq (\leq)\ R_k^o$ for all $k \geq 0$, then $M_k \geq (\leq)\ M_k^o$ for all $k \geq 0$.

The above theorem demonstrates the effects of pessimistic and optimistic designs. If, for example, the upper bounds of noise covariances are known instead of their actual values, the filter may then be designed by setting the noise covariances at their upper bound. This is known as a pessimistic design which gives $M_k \geq M_k^o$ by the above theorem.

Let $\Psi(k,i)$ be the state transition matrix associated with $\bar{F}_k$, i.e.,

$$\Psi(k+1,i) = \bar{F}_k \Psi(k,i) \quad ; \quad \Psi(i,i) = I \quad \forall \quad k \geq i \geq 0 \tag{2.6}$$

From (2.5), M_k^o can then be written as

$$M_k^o = \Psi(k,0)\Pi\Psi'(k,0) + \sum_{i=0}^{k-1} \Psi(k,i+1)[K_i R_i^o K_i' + Q_i^o]\Psi'(k,i+1) \tag{2.7}$$

It can be observed from (2.7) that M_k^o is uniformly bounded (from above), i.e., $\sup_{k\geq 0} \|M_k^o\| < \infty$, if $\bar{F}_k$ is exponentially stable and $[K_k R^o K_k' + Q_k^o]$ is uniformly bounded. Thus, the following theorem which was originally proven by Price (1968) is immediate.

<u>Theorem 2.2</u>: Let F_k, H_k, Q_k, R_k^{-1}, Q_k^o, and R_k^o be uniformly bounded. If M_k is uniformly bounded and $\bar{F}_k$ is exponentially stable, then M_k^o is uniformly bounded.

<u>Remark 2.1</u>: For uniformly bounded F_k, H_k, Q_k, and R_k^{-1}, Anderson and Moore (1981) have shown that M_k is uniformly bounded if $(F_k, R_k^{-\frac{1}{2}} H_k)$ is uniformly detectable, where $R_k = R_k^{\frac{1}{2}}(R_k^{\frac{1}{2}})'$. If, in addition, $(F_k, Q_k^{\frac{1}{2}})$ is uniformly stabilizable, then $\bar{F}_k$ is exponentially stable.

<u>Remark 2.2</u>: For certain high-precision control systems, the uniform boundedness of M_k^o , itself, may not be sufficient for $\hat{x}_{k|k-1}$ to be useful. That is M_k^o may become intolerably large even though it is uniformly bounded. This phenomenon is known as apparent divergence (Fitzgerald, 1971).

For constant F, H, Q, and R, it is well known that under certain conditions (Caines and Mayne, 1970 & 1971; Anderson and Moore, 1979; Chan et al., 1984; De Souza et al., 1986) M_k converges to M which is a solution to the following algebraic Riccati equation (ARE):

$$M = FMF' - FMH'(HMH' + R)^{-1}HMF' + Q \tag{2.8}$$

Let $K := FMH'(HMH' + R)^{-1}$ and $\bar{F} := F - KH$. Then M is called the stabilizing solution of the ARE if $\bar{F}$ has all its eigenvalues inside the unit circle.

Remark 2.3: A necessary and sufficient condition for existence of the stabilizing solution is that (F,H) is detectable and $(F,Q^{1/2})$ has no unreachable mode on the unit circle (Kucera, 1972b; Molinari, 1975; De Souza et al., 1986). It should also be noted that if the stabilizing solution of the ARE exists, then it is necessarily unique.

To the author's knowledge, the convergence of M_k^o, however, has not been rigorously established in any literature. But, it is often used as a part of assumptions in some of literature; see, for example, Mehra (1970) and Toda and Patel (1980). It is clear from (2.5) with constant F, H, Q^o, and R^o that if K_k converges to K and M_k^o converges to M^o, then M^o must satisfy the following algebraic Lyapunov equation (ALE):

$$M^o = \bar{F}M^o\bar{F}' + KR^oK' + Q^o \tag{2.9}$$

In the sequel, linear time-invariant systems with stationary noise processes are considered. Two important properties of the actual one-step predictor error covariance, convergence and divergence, will be studied in Sections 2.2 and 2.3, respectively.

2.2 Convergence Analysis

In this section, we shall establish a sufficient condition for convergence of M_k^o . Cases in which the noise covariances are known up to an unknown scale factor are also considered. For these special cases, it will be shown, in particular, that under certain conditions, the Kalman filter is asymptotically optimal despite incorrect values of noise covariances. Consequently, the residual sequence is asymptotically white under those conditions. This reveals the insufficiency of the whiteness test on the residual sequence used in Mehra (1970). Bounds on the asymptotic filter performance are derived

when the range of errors in noise covariances is known.

Let us first prove the following lemma which will be used for the proof of the convergence of the actual one-step predictor error covariance.

<u>Lemma 2.1</u>: Let A_k be a sequence of square matrices satisfying

(i) $\sup_{k\geq 0} \|A_k\| < \infty$ and

(ii) $\lim_{k\to\infty} A_k = A$ whose all eigenvalues lie inside the unit circle.

Then A_k is exponentially stable, i.e., $\|\Phi(k,i)\| \leq \alpha\beta^{(k-i)}$ $\forall\ k\geq i\geq 0$ for some $\alpha > 0$ and $\beta \in (0,1)$, where $\Phi(k,i)$ denotes the state transition matrix associated with A_k.

<u>Proof</u>: Since eigenvalues of A_k are continuous functions of its entries (Kato, 1976) and A_k converges to A whose all eigenvalues lie inside the unit circle, there exist an integer n_o and $\rho \in (0,1)$ such that all eigenvalues of A_k lie inside or on the circle, centered at the origin, of radius ρ for all $k \geq n_o$, i.e.,

$$\max_j |\lambda_j(A_k)| \leq \rho \quad \forall\ k \geq n_o \tag{2.10}$$

The convergence of A_k also implies that for any $\epsilon > 0$, there exists an integer n_1 such that

$$\sup_{k\geq n_1} \|A_{k+1} - A_k\| \leq \epsilon \tag{2.11}$$

Based on the exponential stability result for slowly varying discrete systems given in Desoer (1970) (see Theorem A.2 in Appendix A), condition (i) together with (2.10) and (2.11) imply that there exist integer n_2, $\gamma > 0$, and $\beta \in (0,1)$ such that

$$\|\Phi(k,i)\| \leq \gamma\beta^{(k-i)} \quad \forall\ k\geq i\geq n_2 \tag{2.12}$$

It then follows from condition (i) and (2.12) that there exists a positive number α such that

$$\|\Phi(k,i)\| \leq \alpha\beta^{(k-i)} \qquad \forall \quad k\geq i\geq 0$$ ∎

Remark 2.4: Suppose that the stabilizing solution M of the **ARE** exists and $\lim_{k\to\infty} M_k = M$. Then, one can observe that $\overline{F}_k := (F - K_k H)$ satisfies both conditions stated in Lemma 2.1. Consequently, $\overline{F}_k$ is exponentially stable. That means if at steady-state conditions the Kalman filter is exponentially stable as a time-invariant filter, then it is also exponentially stable as a time-varying (but asymptotically time-invariant) filter. A somewhat weaker statement of this fact was made without proof by Anderson and Moore (1979, p.82).

The conditions for M_k^o to converge can be established as follows.

Theorem 2.3: Let (F,H) be detectable and $(F,Q^{1/2})$ have no unreachable mode on the unit circle. Suppose that M_k converges to M which is the stabilizing solution of the **ARE**. Then M_k^o converges to M^o which is the unique solution of the **ALE**.

Proof: First notice from Remark 2.3 that the stabilizing solution M of the **ARE** exists and is unique since (F,H) is detectable and $(F,Q^{1/2})$ has no unreachable mode on the unit circle. Here, M_k is uniformly bounded since (F,H) is detectable. Furthermore, $\overline{F}_k$ is exponentially stable as pointed out in Remark 2.4. It then follows from Theorem 2.2 that M_k^o is uniformly bounded. Because of the exponential stability of $\overline{F}$, there exists a unique solution M^o to the **ALE**. But,

$$M_{k+1}^o - M^o = (\overline{F}_k M_k^o \overline{F}_k' + K_k R^o K_k + Q^o) - (\overline{F} M^o \overline{F}' + K R^o K' + Q^o) \qquad (2.13)$$

Since K_k is uniformly bounded and converges to K, K_k may be written as $K_k = K + L_k$ where uniformly bounded matrix sequence L_k converges to zero as k approaches infinity. Thus, (2.13) can be rewritten as

$$M^o_{k+1} - M^o = \bar{F}(M^o_k - M^o)\bar{F}' + \theta_k \tag{2.14}$$

where $\theta_k := - L_k H M^o_k \bar{F}'_k - \bar{F} M^o_k H' L'_k + L_k R^o K'_k + K R^o L'_k$

It is clear that θ_k is uniformly bounded and $\lim_{k\to\infty} \theta_k = 0$. But from (2.14), we have that

$$M^o_k - M^o = \bar{F}^{\,k}(M^o_o - M^o)(\bar{F}')^k + \sum_{i=1}^{k} \bar{F}^{\,k-i}\theta_{i-1}(\bar{F}')^{k-i} \tag{2.15}$$

Since $\bar{F}$ is exponentially stable, the first term on the right hand side (RHS) of (2.15) goes to zero as k approaches infinity. We shall next prove that the second term on RHS of (2.15) also goes to zero as k approaches infinity.

$$\begin{aligned}
\Big\|\sum_{i=1}^{k} \bar{F}^{\,k-i}\theta_{i-1}(\bar{F}')^{k-i}\Big\| &\le \sum_{i=1}^{k} \alpha\beta^{(k-i)}\|\theta_{i-1}\| \quad \text{for some } \alpha > 0\ ,\ \beta \in (0,1) \\
&= \sum_{i=1}^{n} \alpha\beta^{(k-i)}\|\theta_{i-1}\| + \sum_{i=n+1}^{k}\alpha\beta^{(k-i)}\|\theta_{i-1}\| \\
&\le n.\sup_{i\ge 0}\|\theta_i\|\,\alpha\beta^{(k-n)} + \sum_{i=n+1}^{k}\alpha\beta^{(k-i)}\|\theta_{i-1}\|
\end{aligned}$$

Given $\epsilon > 0$. Since $\sum_{i=0}^{\infty}\alpha\beta^i < \infty$ and $\lim_{i\to\infty} \theta_i = 0$, there exist integers N_o and N with $N > N_o$ such that for all $k \ge N$,

$$N_o \sup_{i\ge 0}\|\theta_i\|\,\alpha\beta^{(k-N_o)} \le \epsilon/2 \quad \text{and} \quad \sum_{i=N_o+1}^{k}\alpha\beta^{(k-i)}\|\theta_{i-1}\| \le \epsilon/2.$$

Hence, $\Big\|\sum_{i=1}^{k} \bar{F}^{\,k-i}\theta_{i-1}(\bar{F}')^{k-i}\Big\| \le \epsilon$ for all $k \ge N$.

Consequently, the second term on RHS of (2.15) converges to zero as k approaches infinity. This in turn implies that $\lim_{k\to\infty} M^o_k = M^o$. ■

Remark 2.5: In case that (F,H) is detectable and $(F,Q^{\frac{1}{2}})$ has no unreachable mode on the unit circle, M_k converges to the unique stabilizing solution M of the **ARE** if one of the following conditions holds (Caines and Mayne, 1970 & 1971; Anderson and Moore, 1979; Chan et al., 1984; De Souza et al., 1986):

(i) $(F,Q^{\frac{1}{2}})$ is stabilizable and $M_o \geq 0$.

(ii) $M_o > 0$ or $M_o \geq M$.

Remark 2.6: It should be clear from the proof of Theorem 2.3 that the theorem still holds even though an incorrect initial state error covariance is used, i.e., $M_o \neq \Pi$.

Corollary 2.1: Suppose that F has all its eigenvalues inside the unit circle. Then M_k^o always converges to M^o which is the unique solution of the **ALE**.

Proof: It is clear in this case that (F,H) is detectable and $(F,Q^{\frac{1}{2}})$ is stabilizable. M_k therefore converges to the stabilizing solution of the **ARE**. The convergence of M_k^o to the unique solution of the **ALE** is then immediate from Theorem 2.3. ∎

Remark 2.7: If the filter is designed with no system noise, i.e., $Q = 0$, and the assumption of Corollary 2.1 holds, then the computed error covariance M_k converges to a zero matrix. Consequently, M_k^o converges to M^o which satisfies $M^o = F M^o F' + Q^o$. For this particular case, the limit of M_k^o is therefore independent of R^o.

Let us now define

$$\tilde{y}_{k|k-1} := y_k - H \hat{x}_{k|k-1} \tag{2.16}$$

We shall refer to $\{\tilde{y}_{k|k-1}\}$ as residual sequence. If the filter were designed by using the correct values of noise covariances, $\{\tilde{y}_{k|k-1}\}$

would be a zero-mean white noise sequence (Kailath, 1968). Here, the incorrect values of noise covariances are used, $\tilde{y}_{k|k-1}$ are therefore correlated. Nevertheless, $\tilde{y}_{k|k-1}$ still have mean zero.

For integer h, let us define

$$C(k,h) \quad := \quad \mathrm{COV}(\tilde{y}_{k|k-1},\ \tilde{y}_{k-h|k-h-1}) \quad ; \quad k \geq h \geq 0 \tag{2.17}$$

It is straightforward to show that

$$C(k,h) = \begin{cases} HM^o_kH' + R^o & ; h = 0 \\ H\Psi(k,k-h+1)[\ FM^o_{k-h}H' - K_{k-h}(HM^o_{k-h}H' + R^o)] & ; k \geq h > 0 \end{cases} \tag{2.18}$$

<u>Remark 2.8</u>: It should be observed that if $Q = Q^o$ and $R = R^o$, then $K_{k-h}(HM^o_{k-h}H' + R^o) = FM^o_{k-h}H'$ for $k \geq h > 0$. In such a case, it can be seen from (2.18) that $\{\tilde{y}_{k|k-1}\}$ is white as mentioned earlier. The sequence $\{\tilde{y}_{k|k-1}\}$, in this case, is commonly referred to as innovations sequence.

The following corollary which describes the existence and expression of the limit of C(k,h) is immediate from Theorem 2.3 and (2.18).

<u>Corollary 2.2</u>: Subject to the same conditions given in Theorem 2.3,

$$C_h \ := \ \lim_{k\to\infty} C(k,h) \ = \begin{cases} HM^oH' + R^o & ; h = 0 \\ H\bar{F}^{\,h-1}\ [FM^oH' - K(HM^oH' + R^o)] & ; h > 0 \end{cases} \tag{2.19}$$

<u>Remark 2.9</u>: In case that F has all its eigenvalues inside the unit circle and Q = 0, the expression of C_h can be simplified as

$$C_h \quad = \quad HF^hM^oH' \ + \ R^o\delta_h \quad ; \quad h \geq 0$$

where $M^o = FM^oF' + Q^o$ and δ_h denotes the Kronecker delta, i.e., δ_h is equal to one if h = 0 and equal to zero otherwise.

One can observe from (2.9) and (2.19) that the entries of C_h, Q^o, and R^o are linearly related. If C_h are given and the number of linearly independent equations relating the entries of the actual noise covariances and the entries of C_h is not less than the number of unknown entries of Q^o and R^o, then one can uniquely determine Q^o and R^o. This basic fact has been used in Mehra (1970) and Friedland (1982) to identify the actual noise covariances, in which estimates of C_h are used instead of their actual value in solving for Q^o and R^o.

In some circumstances, the noise covariances are known up to an unknown scale factor (Iglehart and Leondes, 1974; Burg et al., 1982). That is the actual noise covariances are of the form $Q^o = \alpha.Q$ and $R^o = \beta.R$ where α and β are unknown positive scalars. For this special case, the following result can be established.

<u>Proposition 2.1</u>: Subject to the same conditions given in Theorem 2.3, and $Q^o = \alpha.Q$, $R^o = \beta.R$ for some positive scalars α and β , then

$$\text{(i)} \quad M^o = \beta.M + (\alpha - \beta).S_q \tag{2.20}$$

$$\text{and} \quad \text{(ii)} \quad C_h = \begin{cases} \beta.(HMH' + R) + (\alpha - \beta).HS_qH' & ; \quad h = 0 \\ (\alpha - \beta).H\overline{F}^{\,h}S_qH' & ; \quad h > 0 \end{cases} \tag{2.21}$$

where S_q is the unique solution of $\quad S_q = \overline{F}S_q\overline{F}' + Q \qquad (2.22)$

<u>Proof</u>: (i) Since conditions in Theorem 2.3 are satisfied, M_k and M_k^o converge to M and M_o, respectively, and $\overline{F}$ is exponentially stable. From (2.8) and (2.9) together with the exponential stability of $\overline{F}$, we have that

$$M = \sum_{i=0}^{\infty} \overline{F}^{\,i}(KRK' + Q)(\overline{F}')^{i} \quad \text{and} \quad M^o = \sum_{i=0}^{\infty} \overline{F}^{\,i}(\beta.KRK' + \alpha.Q)(\overline{F}')^{i}.$$

Thus, $M^o = \beta.M + (\alpha - \beta).\sum_{i=0}^{\infty} \bar{F}^i Q(\bar{F}')^i$ (2.23)

Because of the exponential stability of $\bar{F}$, $S_q := \sum_{i=0}^{\infty} \bar{F}^i Q(\bar{F}')^i$ is the unique solution of (2.22). Hence, $M^o = \beta.M + (\alpha - \beta).S_q$.

(ii) Substituting $M^o = \beta.M + (\alpha - \beta).S_q$ into (2.19) , one obtains (2.21) after some algebraic manipulations. ∎

An interesting special case arises when $\alpha = \beta$. For this particular case, we obtain the following result.

<u>Corollary 2.3</u>: Subject to the same conditions given in Proposition 2.1 and $\alpha = \beta$, then

(i) The one-step predictor estimate obtained from the Kalman filter using noise covariances Q and R is asymptotically optimal, i.e.,

$$\lim_{k\to\infty} M_k^o = \lim_{k\to\infty} \Sigma_k$$

where Σ_k denotes the minimum, linear one-step predictor error covariance when α is given.

and (ii) The residual sequence is asymptotically white, i.e., $C_h = 0$ for all $h \neq 0$.

<u>Proof</u>: (i) It can be observed that $\alpha^{-1}.\Sigma_k$ and M_k satisfy the same **RDE** with initial conditions $\alpha^{-1}.\Pi$ and Π , respectively. But $\lim_{k\to\infty} M_k = M$ which is the stabilizing solution of the **ARE**. Since the initial values of $\alpha^{-1}.\Sigma_k$ and M_k are different from one another by only a scale factor, it can then be shown (see Theorem B given in Appendix B) that $\lim_{k\to\infty} \Sigma_k = \alpha.M$. From (2.20) with $\alpha = \beta$, $\lim_{k\to\infty} M_k^o = \alpha.M$. The filter is therefore asymptotically optimal.

(ii) With $\alpha = \beta$, it is clear from (2.21) that the residual sequence is asymptotically white. ∎

The above corollary indicates the insufficiency of the whiteness test on the residual sequence, suggested Mehra (1970) to determine whether or not the estimated noise covariances are adequate, in which the filter is assumed to have reached steady-state conditions. It is important to recognize that different noise covariances can lead to the same optimal steady-state gain. But the steady-state covariances of the residual sequences corresponding to those different noise covariances may be quite different. Alspach (1972) pointed out this fact by his numerical examples. Alspach also suggested an additional test which may be performed simultaneously with the whiteness test. The additional test is essentially based on comparing the computed steady-state covariance and the estimated steady-state covariance of the residual sequence, i.e., $HMH' + R$ and $\hat{C}_o$, respectively. As we can see, under the conditions given in Corollary 2.3, the steady-state covariance of the residual sequence C_o is equal to $\alpha.(HMH' + R)$ which is different from $HMH' + R$, unless $\alpha = 1$.

Iglehart and Leondes (1974) proposed several algorithms for estimating α. However, Corollary 2.3 indicates that the filter can be asymptotically optimal although the correct value of α is not known. If the asymptotic filter performance is the only main concern, no estimation of the unknown α is needed under the situation considered in Corollary 2.3. This is a good example to show that better understanding in the behavior of the Kalman filter under incorrect noise covariances can help one avoid unnecessary computations.

In general, $Q = Q^o + \Delta Q$ and $R = R^o + \Delta R$ where ΔQ and ΔR are unknown. Based on Theorem 2.1, bounds on the limit of M_k^o can be easily obtained if ΔQ and ΔR are sign definite of the same sign.

Suppose, for example, that M_k and M_k^o converge to M and M^o, respectively, then $M \geq M^o$ if $\Delta Q \geq 0$ and $\Delta R \geq 0$. The above requirement on sign definiteness of ΔQ and ΔR is, however, restrictive. Instead, one sometimes has information about the range of errors, i.e., $|\Delta q_{ij}|$ and $|\Delta r_{ij}|$ where Δq_{ij} and Δr_{ij} denote the (i,j)th entries of ΔQ and ΔR, respectively. We shall next obtain bounds on the asymptotic filter performance measured by the trace of the limit of M_k^o when the range of errors is known. The analysis given below is similar to the one given by Toda and Patel (1980). Except, we are here working on the actual one-step predictor error covariance instead of the actual filter error covariance, and we are not assuming that F is exponentially stable, as required in Toda and Patel (1980). Let us now introduce the following notations. The Kronecker product of matrices A and B is denoted by $A \otimes B$, and $\mathrm{vec}(A)$ denotes the column vector obtained by concatenating the columns of matrix A. That is $\mathrm{vec}(A) = [\ a_1' \ \ldots \ a_m'\ ]'$ where $\mathbf{a}_i$ denotes the i th column of an $m \times m$ matrix A. The Frobenius norm of A is denoted by $\|A\|_f$, i.e., $\|A\|_f := (\sum_{i,j} |a_{ij}|^2)^{1/2}$. With the above notations, we shall establish the following theorem.

<u>Theorem 2.4</u>: Subject to the same conditions given in Theorem 2.3,

$$\mathrm{tr}(M) - \rho \leq \mathrm{tr}(M^o) \leq \mathrm{tr}(M) + \rho \tag{2.24}$$

where $$\rho := \|K'PK\|_f \cdot \|\Delta R\|_f + \|P\|_f \cdot \|\Delta Q\|_f \tag{2.25}$$

and P is the unique solution to $$P = \bar{F}'P\bar{F} + I \tag{2.26}$$

<u>Proof</u>: Since the conditions in Theorem 2.3 are satisfied, M_k and M_k^o converge to M and M^o, respectively, and $\bar{F}$ is exponentially stable.

But $M^o = \bar{F}M^o\bar{F}' + K(R - \Delta R)K' + (Q - \Delta Q)$

$= \bar{F}M^o\bar{F}' + (KRK' + Q) - (K\Delta RK' + \Delta Q)$

Thus, $\mathrm{vec}(M^o) = \bar{F}\otimes\bar{F}\ \mathrm{vec}(M^o) + \mathrm{vec}(KRK' + Q) - \mathrm{vec}(K\Delta RK' + \Delta Q)$

where we have used the identity (Brewer, 1978) $\mathrm{vec}(ABC) = (C'\otimes A)\mathrm{vec}(B)$.

Let $T := (I_{n^2} - \bar{F}\otimes\bar{F})$, where I_{n^2} denotes an $n^2 \times n^2$ identity matrix and n denotes the dimension of F. Because of the exponential stability of $\bar{F}$, all eigenvalues of $\bar{F}\otimes\bar{F}$ lie inside the unit circle. This in turn implies that T is invertible. Hence,

$$\mathrm{vec}(M^o) = T^{-1}\mathrm{vec}(KRK' + Q) - T^{-1}\mathrm{vec}(K\Delta RK' + \Delta Q) \tag{2.27}$$

Using the identity $\mathrm{tr}(A) = (\mathrm{vec}(I_m))'\mathrm{vec}(A)$ where A is an $m \times m$ matrix, we have that

$$\mathrm{tr}(M^o) = (\mathrm{vec}(I_n))'T^{-1}\mathrm{vec}(KRK' + Q) - (\mathrm{vec}(I_n))'T^{-1}\mathrm{vec}(K\Delta RK' + \Delta Q) \tag{2.28}$$

By performing similar algebraic manipulations, we obtain

$\mathrm{tr}(M) = (\mathrm{vec}(I_n))'T^{-1}\mathrm{vec}(KRK' + Q)$ and $(\mathrm{vec}(P))' = (\mathrm{vec}(I_n))'T^{-1}$.

Thus, (2.28) can be rewritten as

$$\mathrm{tr}(M^o) = \mathrm{tr}(M) - (\mathrm{vec}(P))'\mathrm{vec}(K\Delta RK' + \Delta Q) \tag{2.29}$$

But, $(\mathrm{vec}(P))'\mathrm{vec}(K\Delta RK') = (\mathrm{vec}(K'PK))'\mathrm{vec}(\Delta R)$. Equation (2.29) can then be written as

$$\mathrm{tr}(M^o) = \mathrm{tr}(M) - (\mathrm{vec}(K'PK))'\mathrm{vec}(\Delta R) - (\mathrm{vec}(P))'\mathrm{vec}(\Delta Q) \tag{2.30}$$

Observe that $\|\mathrm{vec}(A)\|_f = \|A\|_f$. Thus,

$$\|(\mathrm{vec}(K'PK))'\mathrm{vec}(\Delta R) + (\mathrm{vec}(P))'\mathrm{vec}(\Delta Q)\|_f \le \|K'PK\|_f\cdot\|\Delta R\|_f + \|P\|_f\cdot\|\Delta Q\|_f$$

It is immediate from (2.30) and the above inequality that (2.24) holds. ■

The significance of the above theorem is that the bounds of the asymptotic filter performance can be computed only by knowledge of the Frobenius norms of ΔQ and ΔR. One should, however, observe that the lower bound of $tr(M^o)$ given in the above theorem is useful only when $tr(M) > \rho$ since $tr(M^o)$ is always nonnegative. It should also be noted that tighter bounds on $tr(M^o)$ than those given above, if desired, may be obtained by using $|\Delta q_{ij}|$ and $|\Delta r_{ij}|$ directly instead of $\|\Delta Q\|_f$ and $\|\Delta R\|_f$.

Remark 2.10: If F has all its eigenvalues inside the unit circle and $Q = 0$, $\lim_{k\to\infty} M_k = M^o$ whose trace is less than or equal to $\|P\|_f \cdot \|\Delta Q\|_f$ where P is the unique solution to $P = F'PF + I$. This is clear since K for this case is a zero matrix.

2.3 Divergence Analysis

So far, we have considered cases in which M_k^o converges. The next subject of our study is the divergence of M_k^o. In particular, we shall show that incorrect values of the system noise covariance can cause the filter to diverge, i.e., the actual one-step predictor error covariance M_k^o becomes unbounded. But if the system is detectable, filter divergence cannot be caused by incorrect values of the measurement noise covariance alone, provided that $R > 0$. For detectable (F,H), one can observe from Theorem 2.3 and Remark 2.5 that if $(F,Q^{1/2})$ is stabilizable, then incorrect values of noise covariances cannot cause the filter to diverge. It is, therefore, necessary that F has at least one unreachable mode associated with an eigenvalue lying outside or on the unit circle if the divergence occurs.

First, let us prove the following two lemmas needed for

establishing the divergence of M_k^o .

Lemma 2.2: Suppose that F has an eigenvalue λ with corresponding left eigenvector z such that $zM_o = zQ = 0$. Then $zM_k = 0 \ \forall\, k \geq 0$.

Proof: Here, $zM_{k+1} = \lambda zM_k[\, F' - H'(HM_kH' + R)^{-1}HM_kF' \,]$.

It then follows immediately from assumption $zM_o = 0$ that $zM_k = 0 \quad \forall\, k \geq 0$. ■

The above lemma simply states that if the mode corresponding to eigenvalue λ of F is assumed to be initially known and not excited by the system noise, i.e., $zM_o = zQ = 0$, then that mode is considered to be exactly known for all time instances, i.e., $zM_k = 0 \ \forall\, k \geq 0$.

Lemma 2.3: Let M satisfy the **ARE**, λ be a complex number with $|\lambda| = 1$, and z be a, possibly complex valued, row vector of appropriate dimension.

Then $z\bar{F} = \lambda z$ if and only if $zF = \lambda z$ and $zQ = 0$. In such a case, necessarily $zK = 0$.

Proof

Sufficiency: Since M satisfies the **ARE**,

$$zMz^* = z[\, FMF' - FMH'(HMH' + R)^{-1}HMF' + Q \,]z^* \tag{2.31}$$

where z^* denotes the complex conjugate transpose of z.

From $zF = \lambda z$ and $zQ = 0$ with $|\lambda| = 1$, (2.31) reduces to

$$zMH'(HMH' + R)^{-1}HMz^* = 0 .$$

Thus, $zK = zFMH'(HMH' + R)^{-1} = 0$. It is then clear that $z\bar{F} = \lambda z$.

Necessity: Since M satisfies the **ARE**,

$$zMz^* = z[\ \bar{F}M\bar{F}' + KRK' + Q\ ]z^* \tag{2.32}$$

But $z\bar{F} = \lambda z$ with $|\lambda| = 1$, (2.32) then reduces to

$$zKRK'z^* + zQz^* = 0 .$$

Hence, $zKRK'z^* = zQz^* = 0$ since both terms are nonnegative. This then implies that $zK = 0$ since R is nonsingular, and $zQ = 0$. Thus, $zF = \lambda z$. ∎

The above lemma simply states that if F has an eigenvalue λ on the unit circle with associated left eigenvector z in the left null space of Q, then $\bar{F}$ has λ as its eigenvalue with the same associated left eigenvector z, and conversely. Let us now give conditions for which M_k^o diverges.

<u>Theorem 2.5</u>: Suppose that F has an eigenvalue λ lying on or outside the unit circle, and its corresponding left eigenvector z is such that $zM_o = zQ = 0$. If $zQ^o \neq 0$, then $zM_k^o z^* \to \infty$ as $k \to \infty$.

<u>Proof</u>: by Lemma 2.2, $zM_k = 0\ \forall\ k \geq 0$. It is then immediate that

$$zK_k = 0 \quad \text{and} \quad z\bar{F}_k = \lambda z \quad \forall\ k \geq 0.$$

But $$M_{k+1}^o = \bar{F}_k M_k^o \bar{F}_k' + K_k R^o K_k' + Q^o .$$

Thus, $$zM_{k+1}^o z^* = |\lambda|^2 . zM_k^o z^* + zQ^o z^* \tag{2.33}$$

Since $|\lambda| \geq 1$ and $zQ^o z^* > 0$, one can then conclude from (2.33) that

$$zM_k^o z^* \to \infty \quad \text{as} \quad k \to \infty .$$ ∎

<u>Remark 2.11</u>: From (2.33), one can observe that $zM_k^o z^*$ will approaches infinity exponentially fast if $|\lambda| > 1$ and linearly fast if $|\lambda| = 1$. Notice also that $zM_k^o z^* = E[\ |z\tilde{x}_{k|k-1}|^2\ ]$. Hence, the above theorem indicates that incorrect values of the system noise covariance

can cause the variance of a certain linear combination of the estimation error to become unbounded.

In case that F has an unreachable mode on the unit circle, it will be seen from the next theorem that M_k can diverge even though the unreachable mode is not assumed to be initially known, i.e., $zM_o \neq 0$.

<u>Theorem 2.6</u>: Subject to (i) M_k converges, and (ii) F has an eigenvalue λ on the unit circle with corresponding left eigenvector z such that $zQ = 0$. If $zQ^o \neq 0$, then $zM_k^oz^* \to \infty$ as $k \to \infty$.

<u>Proof</u>: Suppose that $\lim_{k\to\infty} \inf \; zM_k^oz^* = \alpha$ for some finite nonnegative number α.

From $zM_{k+1}^oz^* = z[\, \bar{F}_kM_k^o\bar{F}_k' + K_kR^oK_k' + Q^o \,]z^*$.

Thus, $$\lim_{k\to\infty} \inf \; zM_{k+1}^oz^* \geq \lim_{k\to\infty} \inf \; z\bar{F}_kM_k^o\bar{F}_k'z^* + \lim_{k\to\infty} \inf \; zK_kR^oK_k'z^* + zQ^oz^* \qquad (2.34)$$

Based on the convergence of M_k and Lemma 2.3, we have that

$$\lim_{k\to\infty} z\bar{F}_k = \lambda z \quad \text{and} \quad \lim_{k\to\infty} zK_k = 0 .$$

Hence, (2.34) reduces to

$$\alpha \geq \alpha + zQ^oz^*$$

Since $zQ^oz^* > 0$, we then reach a contradiction.

Therefore, $\lim_{k\to\infty} \inf \; zM_k^oz^* = \infty$. This in turn implies that

$$zM_k^oz^* \to \infty \quad \text{as} \quad k \to \infty . \qquad \blacksquare$$

<u>Remark 2.12</u>: Notice that conditions (i) and (ii) of Theorem 2.6 can simultaneously hold without assumption $zM_o = 0$. If this were not the case, Theorem 2.6 would be a specialized result of Theorem 2.5. Consider for example:

$$F = 2^{-\frac{1}{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad H = [\,1 \quad 0\,], \; Q = 0, \text{ and } R = 1.$$

Clearly, condition (ii) is satisfied in this case. In addition, (F,H) is observable (and hence detectable), and $M = 0$ is the strong solution of the ARE, i.e., $\bar{F}$ has no eigenvalues outside the unit circle. It then follows from De Souza et al. (1986, Theorem 4.2) that $\lim_{k\to\infty} M_k = 0$ $\forall\, M_o \geq 0$. In particular, if $M_o > 0$, then $\lim_{k\to\infty} M_k = 0$ and $zM_o \neq 0$.

Fitzgerald (1971) investigated divergence of the continuous-time Kalman filter. One of interesting questions addressed by Fitzgerald is whether or not incorrect values of the measurement noise covariance alone can cause the filter to diverge. The answer to this question under restrictive conditions was also given in Fitzgerald (1971). Here, we shall show that filter divergence will not occur for any R, which has been assumed to be positive definite throughout this chapter, if (F,H) is detectable and $Q = Q^o$. That is if the filter is designed with only errors in the measurement noise covariance, then detectability of the system will guarantee uniform boundedness of the actual one-step predictor error covariance. The following lemma, whose proof can be found in Mehra (1976), will be used in establishing the above statement.

Lemma 2.4: Let A and B be symmetric matrices of the same dimension and $B \geq 0$. Then

$$\lambda_{min}(A)\,\mathrm{tr}(B) \quad \leq \quad \mathrm{tr}(AB) \quad \leq \quad \lambda_{max}(A)\,\mathrm{tr}(B)$$

where $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denote the minimum and maximum eigenvalues of A, respectively.

The following theorem can now be established.

<u>Theorem 2.7</u>: Let (F,H) be detectable and $Q = Q^o$. Then M_k^o is uniformly bounded for any $R > 0$.

<u>Proof</u>: Based on (2.4) and (2.5), M_k and M_k^o can be expressed as

$$M_k = \Psi(k,0)\Pi\Psi'(k,0) + \sum_{i=0}^{k-1} \Psi(k,i+1)K_iRK_i'\Psi'(k,i+1) + \sum_{i=0}^{k-1} \Psi(k,i+1)Q\Psi'(k,i+1) \tag{2.35}$$

$$M_k^o = \Psi(k,0)\Pi\Psi'(k,0) + \sum_{i=0}^{k-1} \Psi(k,i+1)K_iR^oK_i'\Psi'(k,i+1) + \sum_{i=0}^{k-1} \Psi(k,i+1)Q^o\Psi'(k,i+1) \tag{2.36}$$

Because of the detectability of (F,H), M_k is uniformly bounded. But each of the terms on the RHS of (2.35) is nonnegative definite. Consequently, each term must be uniformly bounded. From (2.36) with $Q^o = Q$, one can then observe that M_k^o is uniformly bounded if the second term on the the RHS of (2.36) is uniformly bounded. From

$$\mathrm{tr}\Big(\sum_{i=0}^{k-1} \Psi(k,i+1)K_iRK_i'\Psi'(k,i+1) \Big) = \mathrm{tr}\Big(R \sum_{i=0}^{k-1} K_i'\Psi'(k,i+1)\Psi(k,i+1)K_i \Big)$$

$$\geq \lambda_{min}(R) \;\; \xi_k \;\geq\; 0 \quad \text{(by Lemma 2.4)}$$

where $\xi_k := \mathrm{tr}\Big(\sum_{i=0}^{k-1} K_i'\Psi'(k,i+1)\Psi(k,i+1)K_i \Big)$.

Since the second term on the RHS of (2.35) is uniformly bounded, its trace must also be uniformly bounded. In addition, $\lambda_{min}(R) > 0$. Hence, ξ_k is uniformly bounded. Let us denote the induced Euclidean norm of matrix A by $\|A\|_2$, i.e., $\|A\|_2 = (\lambda_{max}(A^*A))^{1/2}$. Then,

$$\Big\| \sum_{i=0}^{k-1} \Psi(k,i+1)K_iR^oK_i'\Psi'(k,i+1)\Big\|_2 = \lambda_{max}\Big(\sum_{i=0}^{k-1} \Psi(k,i+1)K_iR^oK_i'\Psi'(k,i+1) \Big)$$

$$\leq \mathrm{tr}\Big(\sum_{i=0}^{k-1} \Psi(k,i+1)K_iR^oK_i'\Psi'(k,i+1) \Big)$$

$$\leq \lambda_{max}(R^o) \quad \xi_k \qquad \text{(by Lemma 2.4)}$$

The second term on the RHS of (2.36) is therefore uniformly bounded. This in turn implies that M_k^o is uniformly bounded. ∎

2.4 Illustrative Examples

For a purpose of demonstrating the results given Sections 2.2 and 2.3, we consider the following simple scalar system:

$$x_{k+1} = a.x_k + w_k \tag{2.37}$$

$$y_k = x_k + v_k \tag{2.38}$$

where a is a real number and the covariances of x_o, w_k, and v_k are given by $\pi \geq 0$, $q^o > 0$, and $r^o > 0$, respectively.

Suppose that the Kalman filter has been designed with no system noise, i.e., $q = 0$, and measurement noise covariance $r = 1$. The calculated one-step predictor error covariance m_k and the actual one-step predictor error covariance m_k^o satisfy the following difference equations.

$$m_{k+1} = a^2 . m_k / (m_k + 1) \qquad ; m_o = \pi \tag{2.39}$$

$$m_{k+1}^o = b_k . m_k^o + b_k . m_k^2 . r^o + q^o \qquad ; m_o^o = \pi \tag{2.40}$$

where $b_k := a^2 / (m_k + 1)^2$.

Solutions to the above difference equations for $k \geq 0$ can be written as

(i) For $|a| \neq 1$,

$$m_k = c\pi a^{2k} / (c + \pi(a^{2k} - 1)) \tag{2.41}$$

$$m_k^o = \{ c^2\pi + c\pi^2(a^{2k} - 1)r^o + [(c - \pi)^2(1 - a^{-2k})/c$$

$$+ 2\pi(c - \pi)k + \pi^2 a^2 (a^{2k} - 1)/c] \; q^o \} \; a^{2k}/(c + \pi(a^{2k} - 1))^2 \tag{2.42}$$

where $c := a^2 - 1$.

(ii) For $|a| = 1$,

$$m_k = \pi/(\pi k + 1) \tag{2.43}$$

$$m_k^o = (\pi + \pi^2 kr^o + (\pi^2(k+1)(2k+1)/6 + \pi(k+1) + 1)\ kq^o)/(\pi k + 1)^2 \tag{2.44}$$

It is clear that $(F,H) = (a,1)$ is observable (and therefore detectable) for all values of a. Let us now consider the following three possible cases.

Case 1, $|a| < 1$: Here, it is clear that $(F,Q^{\frac{1}{2}}) = (a,0)$ is stabilizable. Thus, m_k converges to the stabilizing solution of the **ARE**, which is zero in this case. It then follows from Theorem 2.3 that m_k^o

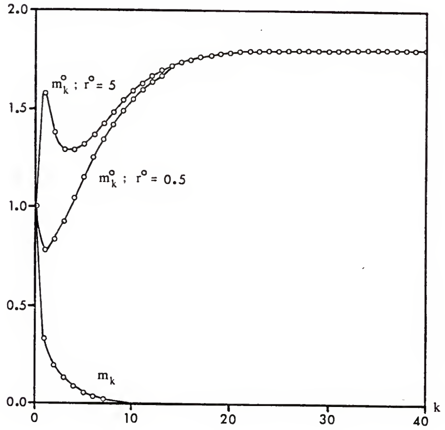

Figure 2.1: Behaviors of m_k and m_k^o for $a = 0.85$, $q^o = 0.5$.

converges to $-q^o/c$ which is the solution of the **ALE**. From (2.41) and (2.42), one can indeed see that $\lim_{k\to\infty} m_k = 0$ and $\lim_{k\to\infty} m_k^o = -q^o/c$ for all $\pi \geq 0$. It should be noted that the limit of m_k^o, in this case, does not depend on the actual measurement noise covariance r^o, as pointed out in Remark 2.7. This fact can also be observed from Fig. 2.1 which depicts behaviors of m_k and m_k^o for $a = 0.85$ with $\pi = 1$ and some values of q^o and r^o.

Case 2, $|a| > 1$: In this case, $(F,Q^{1/2}) = (a,0)$ has no unreachable mode on the unit circle. We shall consider cases $\pi > 0$ and $\pi = 0$, separately.

(i) For $\pi > 0$, condition (ii) of Remark 2.5 is satisfied. Thus, m_k converges to the stabilizing solution of the **ARE** which is equal to c in this case. By Theorem 2.3, m_k^o in turn converges to $(cr^o + a^2q^o/c)$ which is the solution of the **ALE**. The above statements can also be verified by taking the limits of m_k and m_k^o given respectively in (2.41) and (2.42). Behaviors of m_k and m_k^o for $a = 1.5$ with $\pi = 1$ and some values of q^o and r^o are shown in Fig. 2.2. From Fig. 2.2, one can also observe that although $(q^o - q)$ and $(r^o - r)$ are not of the same sign, as required in Theorem 2.1, m_k^o can be less or greater than m_k for all $k > 0$, depending on the values of q^o and r^o.

(ii) For $\pi = 0$, all conditions assumed in Theorem 2.5 are satisfied. Thus, $m_k^o \to \infty$ as $k \to \infty$. One can, in fact, observe from (2.42) that $m_k^o = (a^{2k} - 1)q^o/c$ which approaches infinity exponentially fast. It should be pointed out that since the filter is designed with no system noise and $\pi = 0$, the filter will ignore the measurements y_k completely. The estimate $\hat{x}_{k|k-1}$, in this case, is simply the expectation of x_k.

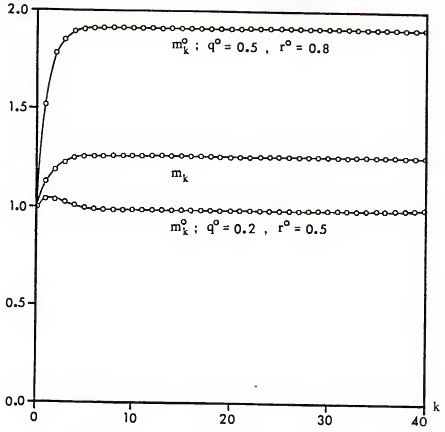

Figure 2.2: Behaviors of m_k and m_k^o for $a = 1.5$.

Case 3, $|a| = 1$: Here, it can be seem from (2.43) that $\lim_{k\to\infty} m_k = 0$ for $\pi > 0$, and $m_k = 0 \;\; \forall\, k \geq 0$ for $\pi = 0$. Based on Theorem 2.6, one then can conclude that $m_k^o \to \infty$ as $k \to \infty$. Using (2.44), one can actually show that m_k^o approaches infinity with the same rate as k. Behaviors of m_k and m_k^o for $|a| = 1$ with $\pi = 1$ and some values of q^o and r^o are depicted in Fig. 2.3. For small q^o (e.g. $q^o = 0.1$), it can be seen from the figure that the actual one-step predictor error covariance does decrease as more measurements are processed for a certain period. But after that period expires, the actual one-step predictor error covariance increases almost linearly as k increases.

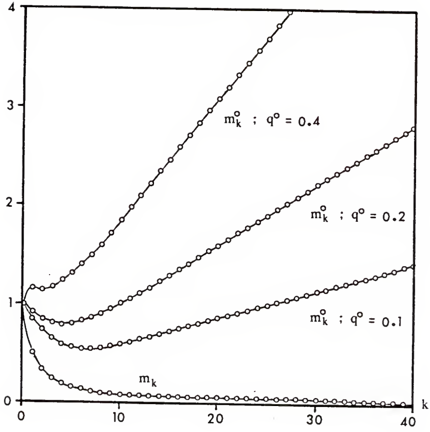

Figure 2.3: Behaviors of m_k and m_k^o for $|a| = 1$, $r^o = 2$.

2.5 Concluding Remarks

Behavior of the discrete-time Kalman filter under incorrect noise covariances has been analyzed in this chapter. The filter performance is quantified by the actual one-step predictor error covariance. Through this quantity, the characteristic of the filter has been studied. Particular emphasis has been given to convergence and divergence properties of the actual one-step predictor error covariance.

Convergence of the actual one-step predictor error covariance, which is often used as a part of assumptions in some of literature, has been established. In addition, it has been shown that under certain

conditions, the Kalman filter is asymptotically optimal even though the noise covariances used in designing the filter are incorrect. The residual sequence is then asymptotically white under those conditions. Bounds on the asymptotic filter performance have been derived when the range of errors in noise covariances is known.

Divergence of the filter has been investigated. It has been shown, in particular, that if the state matrix has an unreachable mode outside or on the unit circle, then incorrect values of the system noise covariance may cause the filter to diverge. On the other hand, if the filter is designed with only errors in the measurement noise covariance and the system is detectable, then filter divergence can never occur.

The contribution of the results presented here is that they help one to understand and be able to predict certain behavior of the Kalman filter when incorrect values of the noise covariances are used. This is, of course, important since the exact values of the noise covariances are hardly known in most practical cases. Continuous-time counterparts of the results presented here will be discussed in the next chapter.

CHAPTER III
BEHAVIOR OF THE CONTINUOUS-TIME KALMAN FILTER UNDER INCORRECT NOISE COVARIANCES

In this chapter, behavior of the continuous-time Kalman filter under incorrect noise covariances is analyzed. The filter performance is quantified by the actual state error covariance. Similar to the previous chapter, Chapter III consists of two main parts, convergence and divergence analyses. In the former part, useful results concerning the convergence of the actual state error covariance are given. Situations in which the actual state error covariance is unbounded are considered in the latter part. In particular, we generalize the divergence results given in Fitzgerald (1971). The results presented in this chapter are based on those given in Sangsuk-Iam and Bullock (1987b).

3.1 Preliminaries

Consider the stochastic dynamical system described by

$$\dot{\mathbf{x}}(t) = F(t)\,\mathbf{x}(t) + \mathbf{w}(t) \tag{3.1}$$

$$\mathbf{y}(t) = H(t)\,\mathbf{x}(t) + \mathbf{v}(t) \tag{3.2}$$

where $\mathbf{x}(t)$ and $\mathbf{y}(t)$ denote the state and the measurement, respectively. It is assumed that system noise process $\{\mathbf{w}(t),\ t\geq 0\}$ and measurement noise process $\{\mathbf{v}(t),\ t\geq 0\}$ are zero-mean white noise processes which are uncorrelated with one another and with $\mathbf{x}(0)$. The autocovariances of $\mathbf{w}(t)$, and $\mathbf{v}(t)$ are given by

$$COV(\mathbf{w}(t),\mathbf{w}(s)) = Q^{o}(t)\delta(t-s) \quad , \quad COV(\mathbf{v}(t),\mathbf{v}(s)) = R^{o}(t)\delta(t-s) \tag{3.3}$$

where $\delta(t)$ denotes the Dirac delta function. For convenience, intensity matrices $Q^o(t)$ and $R^o(t)$ will be referred to as system and measurement noise covariances, respectively.

The linear minimum variance estimate of $x(t)$ given $\{y(s),\ 0\le s\le t\}$ can be obtained from the standard Kalman filter. Let us now consider the case that the filter has been designed with system and measurement noise covariances $Q(t) \ge 0$ and $R(t) > 0$, respectively, instead of the actual noise covariances $Q^o(t)$ and $R^o(t)$. The estimate $\hat{x}(t|t)$ of $x(t)$ given $\{y(s),\ 0\le s\le t\}$ obtained from the filter is then no longer the linear minimum variance estimate. Nevertheless, $\hat{x}(t|t)$ is still an unbiased estimate of $x(t)$.

Let $P(t)$ denote the state error covariance computed from the filter, i.e., $P(t)$ satisfies the following Riccati differential equation (RDE):

$$\dot{P}(t) = F(t)P(t) + P(t)F'(t) - P(t)H'(t)R^{-1}(t)H(t)P(t) + Q(t) \qquad (3.4)$$

where $P(0) = \Pi$ which denotes the covariance of $x(0)$.

Equation (3.4) can also be written in another form as follows.

$$\dot{P}(t) = \bar{F}(t)P(t) + P(t)\bar{F}'(t) + K(t)R(t)K'(t) + Q(t) \qquad (3.5)$$

where $K(t) := P(t)H'(t)R^{-1}(t)$ and $\bar{F}(t) := F(t) - K(t)H(t)$.

The actual state error covariance $P^o(t)$ is defined as the error covariance associated with $\hat{x}(t|t)$,i.e.,

$$P^o(t) := COV(\tilde{x}(t|t),\ \tilde{x}(t|t)) \quad \text{where} \quad \tilde{x}(t|t) := x(t) - \hat{x}(t|t)$$

It can be shown (Nishimura, 1967b) that

$$\dot{P}^o(t) = \bar{F}(t)P^o(t) + P^o(t)\bar{F}'(t) + K(t)R^o(t)K'(t) + Q^o(t) \qquad (3.6)$$

where $P^o(0) = \Pi$.

The following useful result, which can be shown by subtracting (3.6) from (3.5), is due to Nishimura (1967b).

<u>Theorem 3.1</u>: If $Q(t) \geq (\leq)\ Q^o(t)$ and $R(t) \geq (\leq)\ R^o(t)$ for all $t \geq 0$, then $P(t) \geq (\leq)\ P^o(t)$ for all $t \geq 0$.

The effects of pessimistic and optimistic designs can be seen by the above theorem. For instance, if the upper bounds of noise covariances are known instead of the their actual values, the filter may be designed by setting the noise covariances at their upper bound. This is known as a pessimistic design. By the above theorem, the actual state error covariance $P^o(t)$ is therefore bounded from above by the computed state error covariance $P(t)$.

From (3.6), $P^o(t)$ can be expressed as

$$P^o(t) = \Psi(t,0)P^o(0)\Psi'(t,0) + \int_0^t \Psi(t,s)[K(s)R^o(s)K'(s)+Q^o(s)]\Psi'(t,s)ds \qquad (3.7)$$

where $\Psi(t,s)$ denotes the state transition matrix associated with $\overline{F}(t)$, i.e.,

$$\frac{\partial\Psi(t,s)}{\partial t} = \overline{F}(t)\ \Psi(t,s) \quad ; \quad \Psi(s,s) = I \qquad (3.8)$$

One can observe from (3.7) that $P^o(t)$ is uniformly bounded (from above), i.e., $\sup_{t\geq 0} \|P^o(t)\| < \infty$, if $\overline{F}(t)$ is exponentially stable and $[K(t)R^o(t)K'(t)+Q^o(t)]$ is uniformly bounded. This leads to the following theorem which is a slight generalization of a result given by Jazwinski (1970, p.254).

<u>Theorem 3.2</u>: Let $F(t)$, $H(t)$, $Q(t)$, $R^{-1}(t)$, $Q^o(t)$, and $R^o(t)$ be uniformly bounded. If $P(t)$ is uniformly bounded and $\overline{F}(t)$ is exponentially stable, then $P^o(t)$ is uniformly bounded.

Remark 3.1: For uniformly bounded F(t), H(t), Q(t), and $R^{-1}(t)$, a sufficient condition for which P(t) is uniformly bounded is that $(F(t),R^{-\frac{1}{2}}(t)H(t))$ is uniformly completely observable (Anderson, 1971a). If, in addition, $(F(t),Q^{\frac{1}{2}}(t))$ is uniformly completely controllable, then $\overline{F}(t)$ is exponentially stable (Kalman, 1963). It should be noted that uniform detectability and uniform stabilizability for continuous-time systems may be defined in a similar way to the definitions given by Anderson and Moore (1981) for discrete-time systems. The above uniform observability and uniform controllability hypotheses may then be weakened to uniform detectability and uniform stabilizability, respectively.

Remark 3.2: In some practical situations, $P^{o}(t)$ can be intolerably large even though it is uniformly bounded. This phenomenon, which is known as apparent divergence, has been investigated by Fitzgerald (1971).

For time-invariant systems with stationary noise processes, it is well known that under certain conditions (Kwakernaak and Sivan, 1972; Kucera, 1973; Callier and Willems, 1981), P(t) converges to P which is a solution to the following algebraic Riccati equation (ARE):

$$FP + PF' - PH'R^{-1}HP + Q = 0 \tag{3.9}$$

Let $K := PH'R^{-1}$ and $\overline{F} := F - KH$. If P is, in addition, such that $\overline{F}$ has all its eigenvalues with negative real parts, then P is called the stabilizing solution of the ARE.

Remark 3.3: The stabilizing solution P of the ARE exists and is necessarily unique iff (F,H) is detectable and $(F,Q^{\frac{1}{2}})$ has no uncontrollable mode on the imaginary axis (Kucera, 1972a; Molinari, 1973).

The convergence of $P^o(t)$, however, has not been established in any literature, It is often used as a part of assumptions in some of literature, see e.g. Mehra (1970) and Toda and Patel (1978). With constant F,H,Q^o, and R^o, it follows from (3.6) that if $K(t)$ converges to K and $P^o(t)$ converges to P^o, then P^o must satisfy the following algebraic Lyapunov equation (ARE):

$$\overline{F}P^o + P^o\overline{F}' + KR^oK' + Q^o \quad = \quad 0 \tag{3.10}$$

In the following, we shall study certain properties of $P^o(t)$ for linear time-invariant systems with stationary noise processes. Two main properties of $P^o(t)$, convergence and divergence, will be investigated separately in Sections 3.2 and 3.3.

3.2 Convergence Analysis

A sufficient condition for which $P^o(t)$ converges is established in this section. Special cases in which the noise covariances are known up to an unknown scale factor are considered. These special cases are used, in particular, to demonstrate the fact that under certain conditions, the Kalman filter is asymptotically optimal despite incorrect values of noise covariances. The residual process is then asymptotically white under those conditions. Thus, the whiteness test on the residual process suggested by Mehra (1970) is not sufficient to determine whether or not the estimated noise covariances are adequate. Bounds on the asymptotic filter performance measured by the trace of the limit of $P^o(t)$ are also given when the Frobenius norms of ΔQ and ΔR are known.

To prove the convergence of the actual state error covariance $P^o(t)$, we need the following lemma.

Lemma 3.1: Let A(t) be a square matrix whose entries are differentiable functions of t for all $t \in [0, \infty)$. Suppose that A(t) satisfies the following conditions:

(i) $\lim_{t\to\infty} A(t) = A$ whose all eigenvalues have negative real parts, and (ii) $\lim_{t\to\infty} \dot{A}(t) = 0$.

Then A(t) is exponentially stable, i.e., $\|\Phi(t,s)\| \le \alpha e^{-\beta(t-s)}$ for all $t \ge s \ge 0$ for some positive numbers α and β, where $\Phi(t,s)$ denotes the state transition matrix associated with A(t).

Proof: Using continuity and convergence properties of A(t), one can easily show that

$$\sup_{t\ge 0} \|A(t)\| < \infty \tag{3.11}$$

Since eigenvalues of A(t) are continuous functions of its entries (Kato, 1976) and A(t) converges to A whose all eigenvalues have negative real parts, there exist positive numbers t_o and σ such that the real parts of all eigenvalues of A(t) are less than or equal to $-\sigma$ for all $t \ge t_o$, i.e.,

$$\mathrm{Re}\ \lambda_i[A(t)] \le -\sigma \quad \forall\ i \text{ and } t \ge t_o \tag{3.12}$$

The convergence of $\dot{A}(t)$ to a zero matrix implies that for any $\epsilon > 0$, there exist a positive number t_1 such that

$$\sup_{t\ge t_1} \|\dot{A}(t)\| \le \epsilon \tag{3.13}$$

Based on the exponential stability result for slowly time-varying systems proven by Desoer (1969) (see Theorem A.1 given in Appendix A), (3.11) to (3.13) imply that there exist positive numbers t_2, γ, and β such that

$$\|\Phi(t,s)\| \le \gamma e^{-\beta(t-s)} \quad \forall\ t \ge s \ge t_2 \tag{3.14}$$

Since t_2 is finite, there is a positive number ζ such that

$$\|\Phi(t,s)\| \quad \le \quad \zeta \qquad\qquad \forall\ t_2 \ge t \ge s \ge 0\ . \qquad (3.15)$$

It follows directly from (3.14), (3.15), and the semigroup property of the state transition matrix that there exists a positive number α such that

$$\|\Phi(t,s)\| \quad \le \quad \alpha e^{-\beta(t-s)} \qquad \forall\ t \ge s \ge 0\ . \qquad \blacksquare$$

Remark 3.4: It should be noted that the convergence of A(t) to a constant matrix A, in general, does not imply that the limit of $\dot{A}(t)$ exists and equals zero. For example, $a(t) = (\sin t^3)/t \to 0$ as $t \to \infty$. But, $\dot{a}(t) \to \infty$ as $t \to \infty$. However, if in addition the second derivative of A(t) exists and is bounded on $[0,\infty)$, then $\lim_{t\to\infty} \dot{A}(t) = 0$ (Rudin, 1976).

Remark 3.5: If the stabilizing solution P of the **ARE** exists and P(t) converges to P, then $\overline{F}(t) := (F - K(t)H)$ satisfies all of the assumptions stated in Lemma 3.1. Hence, $\overline{F}(t)$ is exponentially stable. That is, if at steady-state conditions the Kalman filter is exponentially stable as a time-invariant filter, then it must be exponentially stable as a time-varying (but asymptotically time-invariant) filter.

A sufficient condition for which $P^o(t)$ converges can now be established as follows.

Theorem 3.3: Let (F,H) be detectable and $(F,Q^{\frac{1}{2}})$ have no uncontrollable mode on the imaginary axis. Suppose that P(t) converges to P which is the stabilizing solution of the **ARE**, then $P^o(t)$ converges to P^o which is the unique solution of the **ALE**.

<u>Proof</u>: Under the above assumptions on F, H, and Q, the stabilizing solution P of the **ARE** exists and is unique. Detectability of (F,H) also implies that P(t) is uniformly bounded. As a consequence of Lemma 3.1 which is stated in Remark 3.5, $\overline{F}(t)$ is exponentially stable. Hence, $P^o(t)$ is uniformly bounded by Theorem 3.2. Because of the exponential stability of $\overline{F}$, there exists a unique solution P^o to the **ALE**.

Let $\Delta(t) := P^o(t) - P^o$. Then,

$$\dot{\Delta}(t) = \overline{F}(t)P^o(t) + P^o(t)\overline{F}'(t) + K(t)R^o(t)K'(t) + Q^o$$
$$- \overline{F}P^o - P^o\overline{F}' - KR^oK' - Q^o \tag{3.16}$$

Since $K(t) := P(t)H'R^{-1}$ is uniformly bounded and converges to K, K(t) can be expressed $K(t) = K + L(t)$ where L(t) is uniformly bounded and converges to a zero matrix as t approaches infinity. Equation (3.16) can then be rewritten as

$$\dot{\Delta}(t) = \overline{F}\Delta(t) + \Delta(t)\overline{F}' + \theta(t) \tag{3.17}$$

where $\theta(t) := -L(t)HP^o(t) - P^o(t)H'L'(t) + L(t)R^oK'(t) + KR^oL'(t)$.

Observe that $\theta(t)$ is uniformly bounded and $\lim_{t\to\infty} \theta(t) = 0$. From (3.17), $\Delta(t)$ can be written as

$$\Delta(t) = e^{\overline{F}t}\Delta(0)\, e^{\overline{F}'t} + \int_0^t e^{\overline{F}(t-s)}\theta(s)\, e^{\overline{F}'(t-s)}ds \tag{3.18}$$

Because of the exponential stability of $\overline{F}$, the first term on the right hand side (RHS) of (3.18) goes to zero as t approaches infinity. Let us now prove that the second term on RHS of (3.18) also goes to zero as t approaches infinity.

$$\left\| \int_0^t e^{\overline{F}(t-s)}\theta(s)\, e^{\overline{F}'(t-s)}ds \right\| \le \int_0^t \alpha e^{-\beta(t-s)}\|\theta(s)\|ds \quad \text{for some } \alpha>0 \text{ and } \beta>0$$

$$= \int_0^\tau \alpha e^{-\beta(t-s)}\|\theta(s)\|ds + \int_\tau^t \alpha e^{-\beta(t-s)}\|\theta(s)\|ds$$

$$\leq \tau\alpha e^{-\beta(t-\tau)} \sup_{s\geq 0} \|\theta(s)\| + \int_{\tau}^{t} \alpha e^{-\beta(t-s)} \|\theta(s)\| ds$$

But $\lim_{t\to\infty} \int_0^t \alpha e^{-\beta(t-s)} ds < \infty$ and $\lim_{t\to\infty} \theta(t) = 0$. It then follows that for any $\epsilon > 0$, there exist positive numbers T_o and T with $T>T_o$ such that for all $t\geq T$,

$$T_o\, \alpha e^{-\beta(t-T_o)} \sup_{s\geq 0}\|\theta(s)\| \leq \epsilon/2 \quad \text{and} \quad \int_{T_o}^{t} \alpha e^{-\beta(t-s)} \|\theta(s)\| ds \leq \epsilon/2.$$

Thus, $\left\| \int_0^t e^{\overline{F}(t-s)} \theta(s)\, e^{\overline{F}'(t-s)} ds \right\| \leq \epsilon$ for all $t \geq T$.

This in turn implies that the second term on RHS of (3.18) goes to zero as t approaches infinity. Consequently, $\lim_{t\to\infty} P^o(t) = P^o$. ■

Remark 3.6: In case that (F,H) is detectable and $(F,Q^{\frac{1}{2}})$ has no uncontrollable mode on the imaginary axis, P(t) converges to the unique stabilizing solution P of the **ARE** if one of the following conditions holds:

(i) $(F,Q^{\frac{1}{2}})$ is stabilizable and $P(0) \geq 0$ (Kwakernaak and Sivan, 1972).

(ii) $P(0) \geq P$ (Pachter and Bullock, 1977; Poubelle et al., 1986).

(iii) $P(0) > 0$ (see Theorem C.1 given in Appendix C).

Remark 3.7: From the proof of Theorem 3.3, it is clear that the theorem still holds even though $P(0) \neq \Pi$, i.e., an incorrect initial state error covariance is used in computing P(t).

Corollary 3.1: Suppose that F has all eigenvalues with negative real parts. Then $P^o(t)$ converges to P^o which is the unique solution of the **ALE**.

Proof: Here, it is obvious that (F,H) is detectable and $(F,Q^{\frac{1}{2}})$ is stabilizable. Consequently, P(t) converges to the stabilizing solution

of the **ARE**. It then follows from Theorem 3.3 that $P^o(t)$ converges to the unique solution of the **ALE**. ■

Remark 3.8: If the assumption of Corollary 3.1 holds and the filter is, in addition, designed by assuming there is no system noise, i.e., $Q = 0$, then the computed state error covariance $P(t)$ converges to a zero matrix. Hence, the actual state error covariance $P^o(t)$ converges to P^o which satisfies $FP^o + P^oF' + Q^o = 0$ and is clearly independent of R^o.

Let

$$\tilde{y}(t|t) \; := \; y(t) - H\,\hat{x}(t|t) \tag{3.19}$$

The process $\{\tilde{y}(t|t),\ t\geq 0\}$ will be referred to as residual process. Let us now define

$$C(t,h) \; := \; COV(\tilde{y}(t|t), \tilde{y}(t-h|t-h)) \quad , \quad t>h\geq 0 \tag{3.20}$$

If the correct values of noise covariances were used in designing the filter, $\{\tilde{y}(t|t),\ t\geq 0\}$ would be a zero-mean white noise process with $C(t,h) = R^o\delta(h)$ (Kailath, 1968). Because of the incorrect values of noise covariances, the residual process is not white. But, the process still has mean zero.

Equation (3.19) can also be written as

$$\tilde{y}(t|t) \; = \; H\,\tilde{x}(t|t) + v(t) \tag{3.21}$$

It is straightforward to show that the estimation error $\tilde{x}(t|t)$ satisfies the following stochastic differential equation.

$$\dot{\tilde{x}}(t|t) \; = \; \bar{F}(t)\,\tilde{x}(t|t) - K(t)\,v(t) + w(t) \tag{3.22}$$

Consequently, $\tilde{x}(t|t)$ can be expressed in terms of stochastic integrals as

$$\tilde{x}(t|t) \; = \; \Psi(t,0)\tilde{x}(0|0) - \int_0^t \Psi(t,s)K(s)v(s)ds + \int_0^t \Psi(t,s)w(s)ds \tag{3.23}$$

Using (3.21) and (3.23), one can show that

$$C(t,h) = \begin{cases} HP^o(t)H' - [HK(t)R^o + R^oK'(t)H']/2 + R^o\delta(0) & ; \ h = 0 \\ H\,\Psi(t,t-h)[P^o(t-h)H' - K(t-h)R^o] \quad + R^o\delta(h) & ; \ t>h>0 \end{cases} \tag{3.24}$$

Remark 3.9: Observe that if $R = R^o$ and $Q = Q^o$, then $K(t-h)R^o$ is equal to $P^o(t-h)H'$ for $t>h\geq 0$. In such a case, one can observe from (3.24) that $C(t,h) = R^o\delta(h)$. The process $\{\tilde{y}(t|t),\ t\geq 0\}$ is therefore white, as pointed out earlier, and it is commonly referred to as innovations process.

Based on Theorem 3.3 and the above equation, the limit of autocovariances $C(t,h)$ of the residual process can be expressed as follows.

Corollary 3.2: Subject to the same conditions given in Theorem 3.3,

$$C_h := \lim_{t\to\infty} C(t,h) = \begin{cases} HP^oH' - (HKR^o + R^oK'H')/2 + R^o\delta(0); & h = 0 \\ H\,e^{\bar{F}h}(P^oH' - KR^o) + R^o\delta(h) & ; \ h > 0 \end{cases} \tag{3.25}$$

It follows from (3.10) and (3.25) that the limiting autocovariances C_h of the residual process have their entries linearly related to the entries of Q^o and R^o. This useful fact was applied by Mehra (1970) to identify the actual noise covariances Q^o and R^o based on the estimates of C_h.

Remark 3.10: The result given in Corollary 3.2 can be further simplified in case that F has all its eigenvalues with negative real parts, and $Q = 0$. for such a case, the limit of $C(t,h)$ always exists and is given by

$$C_h = H\,e^{Fh}P^oH' + R^o\delta(h) \qquad ; \ h \geq 0$$

where $FP^o + P^oF' + Q^o = 0$.

Certain structures of the actual noise covariances may be known in some practical cases (Burg et al., 1982; Iglehart and Leondes, 1974). In particular, the noise covariances may be known up to an unknown scale factor. That is, the actual noise covariances are of the form $Q^o = \alpha.Q$ and $R^o = \beta.R$ where α and β are unknown positive scalars. The following result can be obtained for this particular case.

<u>Proposition 3.1</u>: Subject to the same conditions given in Theorem 3.3, and $Q^o = \alpha.Q$, $R^o = \beta.R$ for some positive scalars α and β, then

$$\text{(i)} \quad P^o = \beta.P + (\alpha - \beta).S_q \tag{3.26}$$

$$\text{(ii)} \quad C_h = (\alpha - \beta).H\, e^{\bar{F}h}\, S_q\, H' + \beta.R\delta(h) \quad ;\ h\geq 0 \tag{3.27}$$

where S_q is the unique solution of
$$\bar{F}S_q + S_q\bar{F}' + Q = 0 \tag{3.28}$$

<u>Proof</u>: (i) Under the conditions given in Theorem 3.3, P(t) and $P^o(t)$ converge respectively to P and P^o, and $\bar{F}$ is exponentially stable. It then follows from (3.9), (3.10), and the exponential stability of $\bar{F}$ that

$$P = \int_0^\infty e^{\bar{F}t}(KRK' + Q)\, e^{\bar{F}'t}dt \tag{3.29}$$

$$P^o = \int_0^\infty e^{\bar{F}t}(\beta.KRK' + \alpha.Q)\, e^{\bar{F}'t}dt \tag{3.30}$$

Hence,
$$P^o = \beta.P + (\alpha - \beta).\int_0^\infty e^{\bar{F}t}\, Q\, e^{\bar{F}'t}dt$$

Because $\bar{F}$ is exponentially stable, the solution of (3.28) is unique and can be written as

$$S_q = \int_0^\infty e^{\bar{F}t}\, Q\, e^{\bar{F}'t}dt$$

Consequently, $P^o = \beta.P + (\alpha - \beta).S_q$.

(ii) Here, $KR^o = \beta.PH'$ and $P^o = \beta.P + (\alpha - \beta).S_q$. Thus,

$$P^oH' - KR^o = (\alpha - \beta).S_qH' \tag{3.31}$$

It follows immediately from (3.25) and (3.31) that (3.27) holds. ■

Let use now consider a special case in which $\alpha = \beta$. Some interesting behavior of the Kalman filter is revealed for this special case. In particular, it will be seen that the Kalman filter can be asymptotically optimal despite incorrect values of α . Thus, as far as the asymptotic filter performance is concerned, the exact value of α is not needed.

<u>Corollary 3.3</u>: Subject to the same conditions given in Proposition 3.1 and $\alpha = \beta$, then

(i) The estimate $\hat{x}(t|t)$ obtained from the Kalman filter using noise covariances Q and R is asymptotically optimal, i.e.,

$$\lim_{t\to\infty} P^o(t) = \lim_{t\to\infty} \Sigma(t)$$

where $\Sigma(t)$ denotes the minimum state error covariance for linear filters when α is given.

and (ii) The residual process is asymptotically white, i.e., $C_h = R^o\delta(h)$.

<u>Proof</u>: (i) We first observe that

$$\dot{\Sigma}(t) = F\Sigma(t) + \Sigma(t)F' - \Sigma(t)H'(\alpha.R)^{-1}H\Sigma(t) + \alpha.Q \quad ; \Sigma(0) = \Pi \tag{3.32}$$

Thus, $\alpha^{-1}.\Sigma(t)$ and $P(t)$ satisfy the same **RDE** with initial values $\alpha^{-1}.\Pi$ and Π , respectively. But $P(t)$ converges to P which is the stabilizing solution of **ARE**. It can be shown (see Theorem C.2 given in Appendix C) that $\alpha^{-1}.\Sigma(t)$ also converges to P since its initial value is different from the initial value of P(t) by only a scale factor. Hence, $\lim_{t\to\infty} \Sigma(t) = \alpha.P$. From (3.26) with $\alpha = \beta$, $\lim_{t\to\infty} P^o(t) = \alpha.P$. Thus,

the filter is asymptotically optimal.

(ii) It follows immediately from (3.27) that the residual process is asymptotically white for $\alpha = \beta$. ∎

In Mehra (1970), a whiteness test on the residual process was suggested to decide whether or not the estimated noise covariances were adequate, in which the filter was assumed to have reached steady-state conditions. It is clear form Corollary 3.3 that the whiteness test is insufficient for the purpose described above. The estimated noise covariances may be quite different from the actual noise covariances although the whiteness test is passed.

It is important to observe that different noise covariances can lead to the same optimal steady-state gain. But, the steady-state autocovariances of the residual processes corresponding to different noise covariances may be quite different. Corollary 3.3 can be used as an example demonstrating the above fact.

In general, $Q = Q^o + \Delta Q$ and $R = R^o + \Delta R$ where ΔQ and ΔR are not completely known. In some cases, one may have information about certain norms of ΔQ and ΔR. Based on this information, certain structures of the limit of $P^o(t)$ can be concluded. The case in which the Frobenius norms of ΔQ and ΔR are known has been investigated by Toda and Patel (1978). In particular, Toda and Patel (1978) have derived the lower and upper bounds on the limit of $tr(P^o(t))$ based on the information about $\|\Delta Q\|_f$ and $\|\Delta R\|_f$, in which F is assumed to be exponentially stable. It is, however, clear from Theorem 3.3 that the limit of $P^o(t)$ may exist even though F is not exponentially stable. Thus, the result proven by Toda and Patel (1978) can be generalized as follows.

<u>Theorem 3.4</u>: Subject to the same conditions given in Theorem 3.3,

$$\mathrm{tr}(P) \; - \; \rho \quad \le \quad \mathrm{tr}(P^o) \quad \le \quad \mathrm{tr}(P) \; + \; \rho \tag{3.33}$$

where $\rho \; := \; \|K'MK\|_f \cdot \|\Delta R\|_f \; + \; \|M\|_f \cdot \|\Delta Q\|_f$ (3.34)

and M is the unique solution to $\bar{F}'M + M\bar{F} \; = \; I$ (3.35)

<u>Proof</u>: Under the conditions given in Theorem 3.3, P(t) and $P^o(t)$ converge, respectively, to P and P^o, and $\bar{F}$ is exponentially stable.

From $\bar{F}P^o + P^o\bar{F}' + K(R - \Delta R)K' + (Q - \Delta Q) \; = \; 0.$

Thus, $\bar{F}P^o + P^o\bar{F}' \; = \; -(KRK' + Q) + (K\Delta RK' + \Delta Q)$ (3.36)

Let $T := (I_n \otimes \bar{F} + \bar{F} \otimes I_n)$, i.e., the Kronecker sum of $\bar{F}$ with itself. Then, (3.36) can be rewritten as

$$T\,\mathrm{vec}(P^o) \quad = \quad -\mathrm{vec}(KRK' + Q) + \mathrm{vec}(K\Delta RK' + \Delta Q) \tag{3.37}$$

Since $\bar{F}$ is exponentially stable, T has all its eigenvalues with negative real parts. This in turn implies that T is invertible. Hence,

$$\mathrm{vec}(P^o) \quad = \quad -T^{-1}\mathrm{vec}(KRK' + Q) + T^{-1}\mathrm{vec}(K\Delta RK' + \Delta Q) \tag{3.38}$$

By performing similar algebraic manipulations to those given in the proof of Theorem 2.4, one can then show that

$$\mathrm{tr}(P^o) \quad = \quad \mathrm{tr}(P) + (\mathrm{vec}(K'MK))'\mathrm{vec}(\Delta R) + (\mathrm{vec}(M))'\mathrm{vec}(\Delta Q) \tag{3.39}$$

Using (3.39) and identity $\|\mathrm{vec}(A)\|_f = \|A\|_f$, one can show that (3.33) holds. ∎

<u>Remark 3.11</u>: It should be noted that the lower bound of $\mathrm{tr}(P^o)$ given in (3.33) is useful only when $\mathrm{tr}(P) > \rho$ since $\mathrm{tr}(P^o)$ is always nonnegative.

<u>Remark 3.12</u>: In case that F has all eigenvalues with negative real parts and $Q = 0$, $P^o(t)$ converges to P^o, the unique solution of the ALE,

whose trace satisfies $tr(P^o) \leq \|M\|_f \cdot \|\Delta Q\|_f$ where M is the unique solution to $F'M + MF = I$. This is due to the fact that K is a zero matrix in this particular case.

3.4 Divergence Analysis

In this section, we investigate situations in which the filter diverges, i.e., the actual state error covariance $P^o(t)$ is unbounded. In particular, we are interested in filter divergence caused by incorrect values of noise covariances. For detectable (F,H), it is clear from Theorem 3.3 and Remark 3.6 that incorrect values of noise covariances cannot cause the filter to diverge if $(F,Q^{1/2})$ is stabilizable. Thus, it is necessary that F has at least one uncontrollable mode associated with an eigenvalue whose real part is nonnegative if the divergence occurs. A special case in which F has uncontrollable modes associated with zero eigenvalues has been previously investigated by Fitzgerald (1971).

Let us now prove the following two lemmas which are useful in establishing conditions for which $P^o(t)$ diverges.

Lemma 3.2: Let z be a row vector with appropriate dimension. If $zP(0) = zQ = 0$, then $zP(t) = 0 \quad \forall\ t \in [0,\infty)$.

Proof: Since P(t) satisfies the RDE,

$$z\dot{P}(0)z^* = z[\ FP(0) + P(0)F' - P(0)H'R^{-1}HP(0) + Q\ ]z^* \quad .$$

where z^* denotes the complex conjugate transpose of z.

It is then clear that $z\dot{P}(0)z^* = 0$. This in turn implies that $zP(t) = 0$ for all $t \in [0,\infty)$ since $zP(0)z^* = 0$. ■

The above lemma simply indicates that if the filter is designed by

assuming that a certain linear combination of the states is initially known and is not excited by the system noise, i.e., $zP(0) = zQ = 0$, then that linear combination of the states is considered to be exactly known for all time instances, i.e., $zP(t) = 0$ for all $t \in [0,\infty)$.

Lemma 3.3: Let P satisfy the **ARE**, λ be a purely imaginary number, and z be a, possibly complex valued, row vector of appropriate dimension.

Then $z\bar{F} = \lambda z$ if and only if $zF = \lambda z$ and $zQ = 0$. In such a case, necessarily $zK = 0$.

Proof

Sufficiency: Since P satisfies the **ARE**,

$$z[\; FP + PF' - PH'R^{-1}HP + Q \;]z^{*} = 0 \tag{3.40}$$

But $zF = \lambda z$ and $zQ = 0$ with $\lambda = j\omega$ for some real number ω. Equation (3.40) then becomes

$$zPH'R^{-1}HPz^{*} = 0 \tag{3.41}$$

Hence, $zK = zPH'R^{-1} = 0$. Consequently, $z\bar{F} = \lambda z$.

Necessity: Since P satisfies the **ARE**,

$$z[\; \bar{F}P + P\bar{F}' + KRK' + Q \;]z^{*} = 0 \tag{3.42}$$

Since $z\bar{F} = \lambda z$ with $\lambda = j\omega$, (3.42) reduces to

$$zKRK'z^{*} + zQz^{*} = 0 \tag{3.43}$$

Thus $zKRK'z^{*} = zQz^{*} = 0$ because both terms are nonnegative. Hence, $zQ = 0$, and $zK = 0$ since $R > 0$. Thus, $zF = \lambda z$. ∎

Remark 3.13: It should be noted that the proof of the above theorem is similar to that of Theorem 1 given in Fitzgerald (1971), except here we are not assuming that $\lambda = 0$.

In words, Lemma 3.3 indicates that if F has an eigenvalue λ on the imaginary axis with associated left eigenvector z in the left null space of Q, then $\bar{F}$ has λ as its eigenvalue with the same associated left eigenvector z, and conversely.

The conditions for which $P^o(t)$ diverges can now be established as follows.

<u>Theorem 3.5</u>: Suppose that F has an eigenvalue λ with nonnegative real part, and its associated left eigenvector z is such that $zP(0) = zQ = 0$. If $zQ^o \neq 0$, then $zP^o(t)z^* \to \infty$ as $t \to \infty$.

<u>Proof</u>: It follows from Lemma 3.2 that $zP(t) = 0$, $\forall\ t \in [0,\infty)$.

Consequently, $zK(t) = zP(t)H'R^{-1} = 0$, $\forall\ t \in [0,\infty)$.

But $z\dot{P}^o(t)z^* = z[\ \bar{F}(t)P^o(t) + P^o(t)\bar{F}'(t) + K(t)R^oK'(t) + Q^o\ ]z^*$

Thus, $$z\dot{P}^o(t)z^* = 2\ \mathrm{Re}\ \lambda . zP^o(t)z^* + zQ^oz^* \tag{3.44}$$

Since $\mathrm{Re}\ \lambda \geq 0$ and $zQ^oz^* > 0$, it is clear form (3.44) that

$$zP^o(t)z^* \to \infty \quad \text{as} \quad t \to \infty .$$ ∎

Observe that $zP^o(t)z^* = E[\ |z\tilde{x}(t|t))|^2]$. The above theorem therefore indicates that the variance of a certain linear combination of the actual estimation error may become unbounded if incorrect values of the system noise covariance are used. It can be seen from (3.44) that $zP^o(t)z^*$ will approaches infinity exponentially fast if $\mathrm{Re}\ \lambda > 0$ and linearly fast if $\mathrm{Re}\ \lambda = 0$ as t approaches infinity.

One of the conditions assumed in Theorem 3.5 is that $zP(0) = 0$. It will be seen from the next theorem that the divergence of $P^o(t)$ can occur even though $zP(0) \neq 0$.

Theorem 3.6: Subject to (i) $P(t)$ converges, and (ii) F has an eigenvalue λ on the imaginary axis with corresponding left eigenvector z such that $zQ = 0$. If $zQ^o \neq 0$, then $zP^o(t)z^* \to \infty$ as $t \to \infty$.

Proof: Suppose that $zP^o(t)z^*$ is finite as t approaches infinity.

From $z\dot{P}^o(t)z^* = z[\ \bar{F}(t)P^o(t) + P^o(t)\bar{F}'(t) + K(t)R^oK'(t) + Q^o\ ]z^*$

Thus
$$\lim_{t\to\infty} \inf z\dot{P}^o(t)z^* \geq \lim_{t\to\infty} \inf z\bar{F}(t)P^o(t)z^* + \lim_{t\to\infty} \inf zP^o(t)\bar{F}'(t)z^*$$
$$+ \lim_{t\to\infty} \inf zK(t)R^oK'(t)z^* + zQ^oz^* \qquad (3.45)$$

If follows from the convergence of $P(t)$ and Lemma 3.3 that

$$\lim_{t\to\infty} z\bar{F}(t) = \lambda z \quad \text{and} \quad \lim_{t\to\infty} zK(t) = 0 \ .$$

It is then clear that the sum of the first three terms on RHS of (3.45) is equal to zero. Hence, $\lim_{t\to\infty} \inf z\dot{P}^o(t)z^* \geq zQ^oz^* > 0$. This then contradicts the hypothesis $zP^o(t)z^*$ is finite as t approaches infinity.

Therefore, $zP^o(t)z^* \to \infty$ as $t \to \infty$. ∎

Remark 3.14: It should be pointed out that conditions (i) and (ii) of the above theorem can simultaneously hold without $zP(0) = 0$. If this were not true, the result of Theorem 3.6 would be included by that of Theorem 3.5. Consider for example:

$$F = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} , \ H = [\ 1 \quad 0\ ] \ , \ Q = 0, \text{ and } R = 1$$

Here, it is clear that condition (ii) holds. Furthermore, (F,H) is detectable, and $P = 0$ is the strong solution of the ARE, i.e., $\bar{F}$ has no eigenvalues with positive real parts. It then follows from Theorem 3 given in Poubelle et al. (1986) that for any $P(0) \geq 0$, $\lim_{t\to\infty} P(t) = 0$. In particular, if $P(0) > 0$, there is no row vector z such that $zP(0) = 0$.

So far, we have shown that incorrect values of the system noise covariance can cause the filter to diverge. The next logical question that we shall investigate is whether or not filter divergence can be caused by the incorrect measurement noise covariance if the correct system noise covariance is used in designing the filter. This question was raised by Fitzgerald (1971). Fitzgerald showed, in particular, that under certain conditions, which were rather restrictive, incorrect values of the measurement noise covariance could not cause the filter to diverge if $Q = Q^o$. Here, a result which is more general than that given by Fitzgerald (1971) will be established. Specifically, it will be shown that if (F,H) is detectable and $Q = Q^o$, then filter divergence will not occur for any R, which has been assumed to be positive definite throughout this chapter. Let us how prove the following theorem.

<u>Theorem 3.7</u>: Let (F,H) be detectable and $Q = Q^o$. Then $P^o(t)$ is uniformly bounded for any $R > 0$.

<u>Proof</u>: From (3.5) and (3.6), P(t) and $P^o(t)$ can be written as

$$P(t) = \Psi(t,0)\Pi\Psi'(t,0) + \int_0^t \Psi(t,s)K(s)RK'(s)\Psi'(t,s)ds + \int_0^t \Psi(t,s)Q\Psi'(t,s)ds \tag{3.46}$$

$$P^o(t) = \Psi(t,0)\Pi\Psi'(t,0) + \int_0^t \Psi(t,s)K(s)R^oK'(s)\Psi'(t,s)ds + \int_0^t \Psi(t,s)Q^o\Psi'(t,s)ds \tag{3.47}$$

Since (F,H) is detectable, P(t) is uniformly bounded. Each individual term on the RHS of (3.46) is nonnegative definite. Thus, each term must be uniformly bounded. It then follows from (3.47) that for $Q = Q^o$, $P^o(t)$ is uniformly bounded if the second term on the RHS of (3.47) is uniformly bounded.

$$\left\| \int_0^t \Psi(t,s)K(s)R^oK'(s)\Psi'(t,s)ds \right\|_2 = \lambda_{max}\left(\int_0^t \Psi(t,s)K(s)R^oK'(s)\Psi'(t,s)ds\right)$$

$$\le tr\left(\int_0^t \Psi(t,s)K(s)R^oK'(s)\Psi'(t,s)ds\right)$$

$$= tr\left(R^o \int_0^t K'(s)\Psi'(t,s)\Psi(t,s)K(s)ds\right)$$

$$\le \lambda_{max}(R^o)\ \xi(t) \quad \text{(by Lemma 2.4)}$$

where $\quad \xi(t) := tr\left(\int_0^t K'(s)\Psi'(t,s)\Psi(t,s)K(s)ds \right)$

We claim that $\xi(t)$ is uniformly bounded. Hence, the second term on the RHS of (3.47) is uniformly bounded. $P^o(t)$ is therefore uniformly bounded. To prove the claim, we observe that

$$tr\left(\int_0^t \Psi(t,s)K(s)RK'(s)\Psi'(t,s)ds \right) \ge \lambda_{min}(R)\ \xi(t) \ge 0 .$$

Since the second term on the RHS of (3.46) is uniformly bounded, its trace must also be uniformly bounded. Furthermore, $\lambda_{min}(R) > 0$. Thus, $\xi(t)$ is uniformly bounded. This completes the proof of the theorem. ∎

3.4 Illustrative Examples

To illustrate the results developed in Sections 3.2 and 3.3, we consider the following scalar system:

$$\dot{x}(t) = a.x(t) + w(t) \tag{3.48}$$

$$y(t) = x(t) + v(t) \tag{3.49}$$

where a is a real number, and the autocovariances of x_o, $w(t)$, and $v(t)$ are given by

$COV(x(0),x(0)) = \pi$, $COV(w(t),w(s)) = q^o\delta(t-s)$, $COV(v(t),v(s)) = r^o\delta(t-s)$ with $\pi \ge 0$, $q^o > 0$, and $r^o > 0$.

Let us now consider the case that the Kalman filter has been designed with no system noise, i.e., $q = 0$, and measurement noise covariance $r = 1$. In this case, the calculated state error covariance $p(t)$ and the actual state error covariance $p^o(t)$ satisfy the following differential equations:

$$\dot{p}(t) = 2a.p(t) - p^2(t) \qquad ; \; p(0) = \pi \qquad (3.50)$$

$$\dot{p}^o(t) = 2(a - p(t))p^o(t) + p^2(t)r^o + q^o \qquad ; \; p^o(0) = \pi \qquad (3.51)$$

Solving the above differential equations, we obtain that for $t \geq 0$:

(i) For $a \neq 0$

$$p(t) = 2a\pi e^{2at}/(2a + \pi(e^{2at} - 1)) \qquad (3.52)$$

$$p^o(t) = \{4a^2\pi + 2a\pi^2(e^{2at} - 1)r^o + [(2a - \pi)^2(1 - e^{-2at})/2a$$

$$+ 2\pi(2a - \pi)t + \pi^2(e^{2at} - 1)/2a]\, q^o\} \; e^{2at}/(2a + \pi(e^{2at} - 1))^2 \qquad (3.53)$$

(ii) For $a = 0$

$$p(t) = \pi/(\pi t + 1) \qquad (3.54)$$

$$p^o(t) = \{\pi + \pi^2 t\, r^o + [(\pi t + 1)^3 - 1] \; q^o/3\pi\} \, /(\pi t + 1)^2 \qquad (3.55)$$

Notice that $(F,H) = (a,1)$ is detectable for all values of a. Let us consider the following three possible cases.

<u>Case 1</u>, $a < 0$: As mentioned in Remark 3.8, $p(t)$ and $p^o(t)$ converge respectively to zero and p^o satisfying $2ap^o + q^o = 0$. Indeed, one can observe from (3.52) and (3.53) that $\lim_{t \to \infty} p(t) = 0$ and $\lim_{t \to \infty} p^o(t) = -q^o/2a$. Behaviors of $p(t)$and $p^o(t)$ for $a = -0.15$ with some values of q^o and r^o are shown in Fig. 3.1. It can also be observed from Fig. 3.1 that the limit of $p^o(t)$, in this case, is independent of the actual measurement noise covariance.

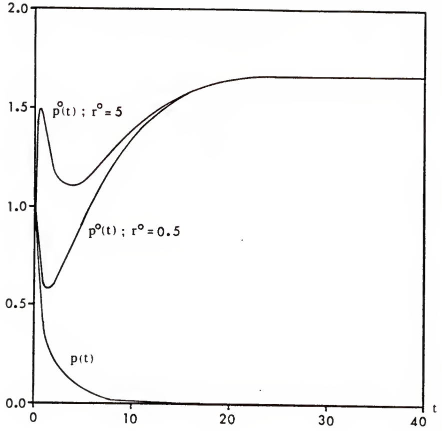

Figure 3.1: Behaviors of $p(t)$ and $p^{o}(t)$ for $a = -0.15$, $q^{o} = 0.5$.

Case 2, $a > 0$: It is clear in this case that $(F, Q^{\frac{1}{2}}) = (a, 0)$ has no uncontrollable mode on the imaginary axis. Let us consider cases $\pi > 0$ and $\pi = 0$ separately.

(i) For $\pi > 0$, $p(t)$ converges to the stabilizing solution of the **ARE** (see condition (iii) of Remark 3.6). Here, the **ARE** is $2ap - p^{2} = 0$. The stabilizing solution is then given by $p = 2a$. By Theorem 3.3, $p^{o}(t)$ converges to p^{o} satisfying $-2ap^{o} + 4a^{2}r^{o} + q^{o} = 0$. From (3.52) and (3.53), one can observe that $\lim_{t\to\infty} p(t) = 2a$ and $\lim_{t\to\infty} p^{o}(t) = 2ar^{o} + q^{o}/2a$ which indeed agree with the above statements. Behaviors of $p(t)$ and $p^{o}(t)$ for $a = 0.15$ with some values of q^{o} and r^{o} are depicted in

Fig. 3.2. Figure 3.2 also shows that $p^o(t)$ can be less or greater than $p(t)$ for all $t > 0$ even though $(q^o - q)$ and $(r^o - r)$ are not of the same sign, as required in Theorem 3.1.

(ii) For $\pi = 0$, all conditions assumed in Theorem 3.5 hold. Therefore, $p^o(t) \to \infty$ as $t \to \infty$. In fact, $p^o(t) = (e^{2at} - 1)q^o/2a$ as obtained from (3.53). Hence, $p^o(t)$ approaches infinity with an exponential rate.

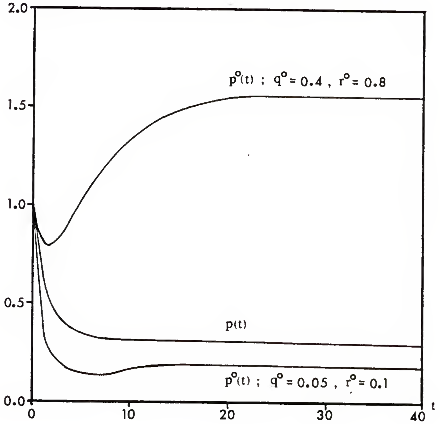

Figure 3.2: Behaviors of $p(t)$ and $p^o(t)$ for $a = 0.15$.

Case 3, $a = 0$: It is clear from (3.54) that $\lim_{t\to\infty} p(t) = 0$ for $\pi > 0$ and $p(t) = 0 \ \forall \ t \geq 0$ for $\pi = 0$. Thus, all conditions assumed in Theorem 3.6 are satisfied. This in turn implies that $p^o(t) \to \infty$ as $t \to \infty$. In

fact, one can observe from (3.55) that $p^o(t)$ approaches infinity with the same rate as t. Behaviors of p(t) and $p^o(t)$ with $\pi = 1$, $r^o = 2$ and several values of q^o are shown in Fig. 3.3. It can be seen from Fig. 3.3 that for small q^o , e.g., $q^o = 0.1$, the actual state error covariance $p^o(t)$ does decrease as more measurements are processed for a short period. But for large t, $p^o(t)$ increases almost linearly as t increases.

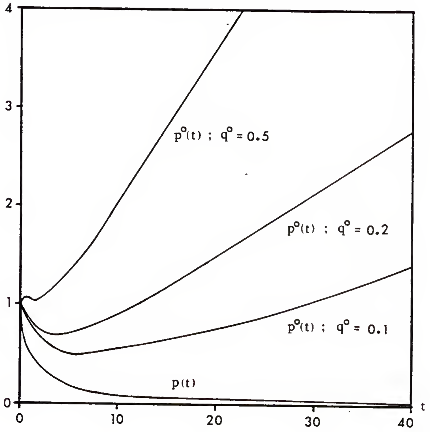

Figure 3.3: Behaviors of p(t) and $p^o(t)$ for a = 0, $r^o = 2$.

3.5 Concluding Remarks

In this chapter, we have analyzed behavior of the continuous-time Kalman filter under incorrect noise covariances. The actual state error covariance is used as a quantity representing the filter performance.

Based on this quantity, the characteristic of the filter has been examined. In particular, two important properties of the actual state error covariance, convergence and divergence, have been investigated.

A sufficient condition for convergence of the actual state error covariance has been established. It is shown, in addition, that under certain conditions the Kalman filter is asymptotically optimal despite incorrect values of noise covariances. Under these conditions, the residual process is then asymptotically white.

Situations in which the filter diverges has been investigated. In particular, it has been shown that incorrect values of the system noise covariance may cause the filter to diverge if the state matrix has uncontrollable modes associated with eigenvalues whose real parts are nonnegative. But if the system is detectable, incorrect values of the measurement noise covariance alone cannot cause the filter to diverge.

It can be seen that the discrete-time and continuous-time results presented respectively in the previous chapter and this chapter are very similar. In most cases, one can immediately write down what the corresponding continuous-time results would be if one knows the discrete-time results, and vice versa.

CHAPTER IV
DIRECT ESTIMATION OF NOISE COVARIANCES

In some practical cases, the noise covariances are unknown. For such cases, one might want to estimate the unknown noise covariances and use the estimated noise covariances in designing the Kalman filter. Numerous techniques for estimating noise covariances have been proposed, see for example Mehra (1970, 1972), Weiss (1970), Belanger (1974), Myers and Tapley (1976), Tabuchi et al. (1978), Ohnishi (1980), Lee (1980), Friedland (1982) and Lee et al. (1985). For linear time-invariant systems with stationary noise processes, most of commonly used techniques, e.g., measurement correlation, residual sequence correlation, and covariance matching techniques, have at least one of the following drawbacks. First, the estimation of the state is required as an intermediate step. Second, the measurements or the residual sequence must reach its stationary stage before one starts estimating the unknown noise covariances.

In Lee (1980), an interesting and rather direct technique was suggested. By utilizing the minimal polynomial of the state matrix, Lee showed that a certain linear combination of the measurements was stationary although the measurements themselves were not stationary. Based on the sample autocovariances of that linear combination of the measurements, the unknown noise covariances can be estimated. Clearly, this technique does not possess the undesirable drawbacks mentioned above. It is, however, unfortunate that Lee merely introduced his idea

and did not give any analyses on his technique. Particularly, convergence of the estimated values of the noise covariances to their actual values was not discussed.

In this chapter, we shall discuss and analyze the direct technique suggested by Lee (1980) which will be referred to as the stationary preprocessed measurement correlation technique. The technique is direct in the sense that the unknown noise covariances can be estimated without requiring the estimate of the state and the stationarity of the measurements.

4.1 Stationary Preprocessed Measurements

Consider the following discrete stochastic system

$$x_{k+1} = F\,x_k + w_k \tag{4.1}$$

$$y_k = H\,x_k + v_k \tag{4.2}$$

where w_k and v_k are zero-mean independent noise processes with unknown covariances Q and R, respectively. It is also assumed that noise processes w_k and v_k are mutually independent.

Given measurements $\{y_k,\ 0 \le k \le N\}$, we would like to identify the unknown noise covariances. The identification can be achieved by extracting the necessary information about the unknown noise covariances from the measurements. Because the measurements are, in general, nonstationary, difficulty in extracting the necessary information arises. This difficulty can, however, be circumvented by preprocessing the measurements to remove nonstationary components. We shall refer to these preprocessed measurements as stationary preprocessed measurements. It should be pointed out that the preprocessing technique mentioned above is commonly used in time series analysis; see Anderson (1971),

Box and Jenkins (1976), and Aoki (1987). For example, polynomial trends of a given time series can be removed after a certain number of differencing. Taking difference of the logarithms of the time series, on the other hand, can remove exponential growth trends.

Two different approaches for obtaining stationary preprocessed measurements have been proposed by Ohnishi (1980) and Lee (1980). The approach proposed by Ohnishi is, however, unnecessarily complicated. Futhermore, the dimension of the stationary preprocessed measurements obtained from Ohnishi's approach is equal to the dimension of the state which is usually greater than the dimension of the original measurements. Lee's approach, on the other hand, is much simpler. By utilizing the minimal polynomial of the state matrix F, Lee showed that a certain linear combination of the measurements was stationary even though the measurements themselves were not stationary. That linear combination of the measurements was then used as a stationary preprocessed measurement. Clearly, the dimension of the stationary preprocessed measurements obtained from Lee's approach remains the same as the dimension of the original measurements.

Let us now show how the stationary preprocessed measurements can be obtained by utilization of the minimal polynomial of F. Let the minimal polynomial of F be given by

$$f(d) = d^m + a_1 d^{m-1} + \ldots + a_m \tag{4.3}$$

From (4.1) and (4.2), we have that

$$y_{k-m} = H x_{k-m} + v_{k-m}$$

$$y_{k-m+1} = HF x_{k-m} + H w_{k-m} + v_{k-m+1}$$

$$y_{k-m+2} = HF^2 x_{k-m} + HF\, w_{k-m} + H\, w_{k-m+1} + v_{k-m+2}$$

$$\vdots \qquad \vdots$$

$$y_k = HF^m x_{k-m} + \sum_{j=1}^{m} HF^{j-1} w_{k-j} + v_k$$

Define $z_k := y_k + a_1\, y_{k-1} + \ldots + a_m\, y_{k-m}\ ,\ k \geq m$ (4.4)

Since $f(F) = 0$, it follows that

$$z_k = \sum_{i=1}^{m} (HA_i\, w_{k-i} + a_i\, v_{k-i}) + v_k \tag{4.5}$$

where $A_i := \sum_{j=0}^{i-1} a_j\, F^{i-j-1}$, $i = 1,2,..,m$ with $a_o := 1$. (4.6)

Notice that A_i can be computed recursively as

$$A_{i+1} = FA_i + a_i I \quad \text{for } i = 1,2,..,m-1 \quad \text{with } A_1 = I.$$

It is apparent from (4.5) that z_k is (wide-sense) stationary. Thus, stationary preprocessed measurement z_k is obtained.

4.2 Estimation of Noise Covariances

In this section, we describe how one can utilize the correlation of stationary preprocessed measurements z_k to obtain the estimates of the unknown noise covariances. It is clear from (4.5) that the process z_k is not only stationary but also m-dependent, i.e., z_i and z_j are independent if $|i-j| > m$ (further discussions on m-dependent time series can be found in Hoeffding and Robbins (1948) and Fuller (1976)). Hence, z_k has zero autocovariances after time lag m.

Let $C_h := \mathrm{COV}(\, z_k,\ z_{k-h}\,)$ (4.7)

Then
$$C_o = H(A_1 Q\, A_1' + A_2 Q\, A_2' + \ldots + A_m Q\, A_m')H' + (1 + a_1^2 + \ldots + a_m^2\,)\, R$$

$$C_1 = H(A_2 Q A_1' + A_3 Q A_2' + \ldots + A_m Q A_{m-1}')H' + (a_1 + a_2 a_1 + \ldots + a_m a_{m-1}) R$$

$$\vdots \qquad \vdots$$

$$C_i = H(A_{i+1} Q A_1' + A_{i+2} Q A_2' + \ldots + A_m Q A_{m-i}')H' + (a_i + a_{i+1} a_1 + \ldots + a_m a_{m-i}) R$$

$$\vdots \qquad \vdots$$

$$C_m = a_m R \tag{4.8}$$

From the above expressions of C_h, one can see that the entries of C_h, Q, and R are linearly related. Hence, a set of linear equations which may be used to solve for Q and R can be obtained if C_o, $C_1, \ldots, C_m$ are given. In particular, $R = a_m^{-1} . C_m$ provided that a_m is not zero, or equivalently F is invertible (Herstein, 1975).

Given measurements $\{y_k \;,\; 0 \le k \le N\}$, the autocovariances C_h of z_k may be estimated by

$$\hat{C}_h = \Big(\sum_{i=m+h}^{N} z_i z_{i-h}'\Big)/(N-m-h+1) \qquad , h = 0,1,\ldots,m \tag{4.9}$$

The above equation can be written in a recursive form as

$$\hat{C}_h(N) = \frac{(N-m-h)}{(N-m-h+1)} \hat{C}_h(N-1) + \frac{1}{(N-m-h+1)} z_N z_{N-h}' \;,\; N \ge m+h \tag{4.10}$$

where $\hat{C}_h(j)$ denotes the estimate of C_h based on measurements $\{y_k \;,\; 0 \le k \le j\}$.

Using the stationary property of z_k, one can easily show that $\hat{C}_h$ is an unbiased estimate of C_h. If the number of linearly independent equations relating the entries of C_h, Q and R is not less than the number of unknown entries of Q and R, the estimates of Q and R can be

uniquely determined from $\hat{C}_h$. We shall refer to the technique for estimating the noise covariances described above as the stationary preprocessed measurement correlation (SPMC) technique. The convergence property of the SPMC technique will be discussed in the next section.

Let us now make an observation that a linear combination of the measurements possessing stationarity can be defined by using any polynomial p(d) which is satisfied by the state matrix, i.e., p(F) = 0. As well known, it is in general difficult to obtain the minimal polynomial of a given matrix. Therefore, it is worthwhile to investigate whether there is benefit in defining z_k based on the minimal polynomial of F.

Let p(d) be a polynomial which is satisfied by F and n(p) be the number of independent equations, with respect to entries of Q and R, which are obtained from autocovariance equations (4.8) of stationary preprocessed measurement $\bar{z}_k$ defined as in (4.4) by using polynomial p(d) instead. The benefit of defining stationary preprocessed measurements based on the minimal polynomial of F is revealed by the following proposition.

Proposition 4.1: Let f(d) be the minimal polynomial of F. Then for any polynomial p(d) which is satisfied by F, $n(f) \geq n(p)$.

Proof: Since f(d) is the minimal polynomial of F, p(d) can be factored as

$$p(d) = g(d).f(d) \tag{4.11}$$

for some polynomial $g(d) = \sum_{i=0}^{s} b_i d^{s-i}$.

Let z_k and $\bar{z}_k$ be linear combinations of y_k defined as in (4.4) by using polynomials f(d) and p(d), respectively.

From (4.11), $\bar{z}_k$ can be expressed in terms of z_k as

$$\bar{z}_k = \sum_{i=0}^{s} b_i \, z_{k-i} \tag{4.12}$$

Let $\bar{C}_h$ and C_h be defined as $COV(\bar{z}_k, \bar{z}_{k-h})$ and $COV(z_k, z_{k-h})$, respectively.

Then $$\bar{C}_h = COV(\sum_{i=0}^{s} b_i \, z_{k-i} , \sum_{j=0}^{s} b_j \, z_{k-h-j})$$

$$= \sum_{i=0}^{s} \sum_{j=0}^{s} b_i \, b_j \, C_{h+j-i} \tag{4.13}$$

That is $\bar{C}_h$ can be written as a linear combination of C_h. This in turn implies that $n(f) \geq n(p)$. ∎

4.3 Convergence of Noise Covariance Estimates

In this section, we investigate the convergence property of the noise covariance estimates obtained from the SPMC technique. Since the estimate values of noise covariances are linear functions of $\hat{C}_h$, their convergence to the actual values depends thoroughly on the convergence of $\hat{C}_h$ to C_h. As mentioned earlier, for non-zero a_m the measurement noise covariance can be estimated by

$$\hat{R} = a_m^{-1} \cdot \hat{C}_m \tag{4.14}$$

But if $\hat{C}_m$ does not converge to C_m, $\hat{R}$ may not be even a symmetric positive semi-definite matrix. For such a case, the estimate of R is not useful. Hence, the establishment of the convergence of $\hat{C}_h$ is essential.

Let us now prove he following theorem which can be directly applied to establish the convergence of $\hat{C}_h$.

<u>Theorem 4.1</u>: Let $\{u_k, \ k = 0,1,..\}$ be a stationary and m-dependent sequence of zero-mean random vectors with $\sup_{k,j} E[u_{k,j}^4] < \infty$, where $u_{k,j}$ denotes the j th component of u_k. Then

$$\frac{1}{N-h} \sum_{i=h}^{N-1} u_i \, u_{i-h}' \xrightarrow[N \to \infty]{} R_h \qquad \text{in quadratic mean}$$

where $R_h \;:=\; \mathrm{COV}(u_i, u_{i-h})$.

Proof: Without loss of generality, we shall prove the theorem by using the induced Euclidean norm.

Here, $\left\| \frac{1}{N-h} \sum_{i=h}^{N-1} u_i\, u_{i-h}' - R_h \right\| := \sup_{\|v\|=1} \left\| \left(\frac{1}{N-h} \sum_{i=h}^{N-1} u_i\, u_{i-h}' - R_h \right) v \right\|$

Where $\|v\|^2 \;:=\; v'v$.

To prove $\lim_{N\to\infty} E\left[\left\| \frac{1}{N-h} \sum_{i=h}^{N-1} u_i\, u_{i-h}' - R_h \right\|^2 \right] = 0$, it is sufficient show that for any v with $\|v\| = 1$,

$$\lim_{N\to\infty} E\left[\left\| \left(\frac{1}{N-h} \sum_{i=h}^{N-1} u_i\, u_{i-h}' - R_h \right) v \right\|^2 \right] = 0 \tag{4.15}$$

From $E\left[\left\| \left(\frac{1}{N-h} \sum_{i=h}^{N-1} u_i\, u_{i-h}' - R_h \right) v \right\|^2 \right]$

$$= \frac{1}{(N-h)^2} \sum_{i=h}^{N-1} \sum_{j=h}^{N-1} v'\, E[(u_i\, u_{i-h}' - R_h)'(u_j\, u_{j-h}' - R_h)]\, v$$

$$= \frac{1}{(N-h)^2} \sum_{i=h}^{N-1} \sum_{k=i-N+1}^{i-h} v'\, E[(u_i\, u_{i-h}' - R_h)'(u_{i-k}\, u_{i-k-h}' - R_h)]\, v$$

$$= \frac{1}{(N-h)^2} \sum_{k=h-N+1}^{0} \sum_{i=h}^{N-1+k} v'\, E[(u_i\, u_{i-h}' - R_h)'(u_{i-k}\, u_{i-k-h}' - R_h)]\, v$$

$$+ \frac{1}{(N-h)^2} \sum_{k=1}^{N-1-h} \sum_{i=k+h}^{N-1} v'\, E[(u_i\, u_{i-h}' - R_h)'(u_{i-k}\, u_{i-k-h}' - R_h)]\, v$$

Now observe that $u_i\, u_{i-h}'$ and $u_{i-k}\, u_{i-k-h}'$ are idependent for $|k| > m+h$. Thus

$$E[(u_i\, u_{i-h}' - R_h)'(u_{i-k}\, u_{i-k-h}' - R_h)] = 0 \quad \text{for } |k| > m+h .$$

Thus, for $N > m+2h$

$E\left[\left\| \left(\frac{1}{N-h} \sum_{i=h}^{N-1} u_i\, u_{i-h}' - R_h \right) v \right\|^2 \right]$

$$= \frac{1}{(N-h)^2} \sum_{k=-m-h}^{0} \sum_{i=h}^{N-1+k} v'\, E[(u_i\, u_{i-h}' - R_h)'(u_{i-k}\, u_{i-k-h}' - R_h)]\, v$$

$$+ \frac{1}{(N-h)^2} \sum_{k=1}^{m+h} \sum_{i=k+h}^{N-1} v'\, E[(u_i\, u_{i-h}' - R_h)'(u_{i-k}\, u_{i-k-h}' - R_h)]\, v$$

$$\leq \frac{1}{(N-h)^2} (m+h+1) \sum_{i=h}^{N-1} \max_{-m-h \leq k \leq 0} |v' E[(u_i u'_{i-h} - R_h)'(u_{i-k} u'_{i-k-h} - R_h)] v|$$

$$+ \frac{1}{(N-h)^2} (m+h) \sum_{i=h}^{N-1} \max_{1 \leq k \leq m+h} |v' E[(u_i u'_{i-h} - R_h)'(u_{i-k} u'_{i-k-h} - R_h)] v|$$

But $E[(u_i u'_{i-h} - R_h)'(u_{i-k} u'_{i-k-h} - R_h)] = E[u_{i-h} u'_i u_{i-k} u'_{i-k-h}] - R'_h R_h$. Based on the above equation and $\sup_{k,j} E[u^4_{k,j}] < \infty$, one can show that each entry of $E[(u_i u'_{i-h} - R_h)'(u_{i-k} u'_{i-k-h} - R_h)]$ is finite for all h, k, and i. It is then immediate that

$$E[\|(\frac{1}{N-h} \sum_{i=h}^{N-1} u_i u'_{i-h} - R_h)v\|^2] \leq \frac{c}{(N-h)} \quad \text{for some finite number c.}$$

Thus, equation (4.15) holds. This in turn implies that

$$\frac{1}{N-h} \sum_{i=h}^{N-1} u_i u'_{i-h} \xrightarrow[N \to \infty]{} R_h \quad \text{in quadratic mean.} \qquad \blacksquare$$

It is clear from the above theorem that the sample autocovariances $\hat{C}_h$ will converge in quadratic mean to the actual autocovariances C_h as N increases if $\sup_{k,j} E[w^4_{k,j}]$ and $\sup_{k,j} E[v^4_{k,j}]$ are finite. Under the same conditions, the estimates of the noise covariances will also converge in quadratic mean to the actual values as N increases since the estimates are linear functions of $\hat{C}_h$. It should also be noted that convergence in quadratic mean implies convergence in probability (Chung, 1974; Wong and Hajek, 1985). Thus , under the above conditions on the fourth moments of the noise processes, the estimates of the noise covariances obtained from the SPMC technique are not only unbiased but also consistent.

4.4 Unique Determinability of Noise Covariances

Given C_h, the noise covariances Q and R can be uniquely determined if the number of unknown entries of Q and R is less than or equal to

$n(f)$. It is, however, important to observe that the measurement noise covariance R can always be uniquely determined, provided that the state matrix F is invertible. It should also be pointed out that when F is invertible, H does not have to be known if one wants to determine only R. In case that F is singular, one may or may not be able to determine R uniquely. The following example demonstrates the preceding statement.

<u>Example 4.1</u>: Consider the system described by (4.1) and (4.2) with

$$F = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \qquad Q = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}$$

Here, $f(d) = d^3 + a_1 d^2 + a_2 d + a_3$ where $a_1 = -2$, $a_2 = 1$, and $a_3 = 0$.

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} -2 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$C_o = H(A_1QA_1' + A_2QA_2' + A_3QA_3')H' + (1 + a_1^2 + a_2^2 + a_3^2)R$$

$$= \begin{bmatrix} 6q_1 + 2q_2 + 2q_3 + 6r_1 & -q_3 \\ -q_3 & 2q_3 + 6r_2 \end{bmatrix} \tag{4.16}$$

$$C_1 = H(A_2QA_1' + A_3QA_2')H' + (a_1 + a_2a_1 + a_3a_2)R$$

$$= \begin{bmatrix} -4q_1 - q_2 + q_3 - 4r_1 & 0 \\ 0 & -q_3 - 4r_2 \end{bmatrix} \tag{4.17}$$

$$C_2 = H(A_3QA_1')H' + (a_2 + a_1a_3)R$$

$$= \begin{bmatrix} q_1 + r_1 & q_3 \\ 0 & r_2 \end{bmatrix} \tag{4.18}$$

Given C_o, C_1, and C_2, it can be observed from (4.16)–(4.18) that only q_2, q_3, r_2, and $(q_1 + r_1)$ can be uniquely determined. Thus R and also Q cannot be uniquely determined in this case. On the other hand, if it is known beforehand that the first component of the state is not directly excited by the system noise, i.e., $q_1 = 0$, the measurement noise covariance R can be uniquely determined.

In case that the number of unknown entries of Q and R is greater than n(f), the SPMC technique does not provide unique estimates of the noise covariances. For such a case, one might want to estimate the optimal filter steady-state gain, if it exists, instead. There are several techniques available for estimating the optimal filter steady-state gain, see Mehra (1970), Carew and Belanger (1973), Neethling and Young (1974), Tse and Weinert (1975), and Tajima (1978).

4.5 Simulation Results

A simulation study of the SPMC technique is given in this section. The system used in the simulation study is described by (4.1) and (4.2) with

$$F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad H = [\ 1 \quad 0\ ], \quad Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}, \quad R = r$$

and $\mathbf{w}_k$ and $\mathbf{v}_k$ are Gaussian white noise sequences.

The minimal polynomial of F is given by

$$f(d) = d^2 - 2d + 1$$

Thus, the stationary preprocessed measurement can be expressed as

$$z_k = y_k - 2y_{k-1} + y_{k-2}$$

Here, the autocovariances of z_k are

$$C_o = 2q_1 + q_2 + 6r$$

$$C_1 = -q_1 - 4r$$

$$C_2 = r$$

Given measurements of $\{y_k,\ 0 \le k \le N\}$, the estimates of C_o, C_1, and C_2 can be obtained according to (4.9). The estimates of q_1 , q_2, and r can then be given by

$$\hat{r} = \hat{C}_2$$

$$\hat{q}_1 = -(\hat{C}_1 + 4\hat{r})$$

$$\hat{q}_2 = \hat{C}_o - 2\hat{q}_1 - 6\hat{r}$$

NOISE VARIANCES	q_1	q_2	r
ACTUAL VALUE	1.5	1.0	0.5
RUN	ESTIMATES		
NUMBER	$\hat{q}_1$	$\hat{q}_2$	$\hat{r}$
1	1.503	1.002	0.489
2	1.790	1.021	0.437
3	1.328	1.160	0.496
4	1.555	0.813	0.446
5	1.538	1.001	0.506
6	1.677	0.939	0.461
7	1.489	0.933	0.534
8	1.420	1.053	0.543
9	1.550	1.094	0.467
10	1.509	0.982	0.485

Table 4.1: Simulation Results.

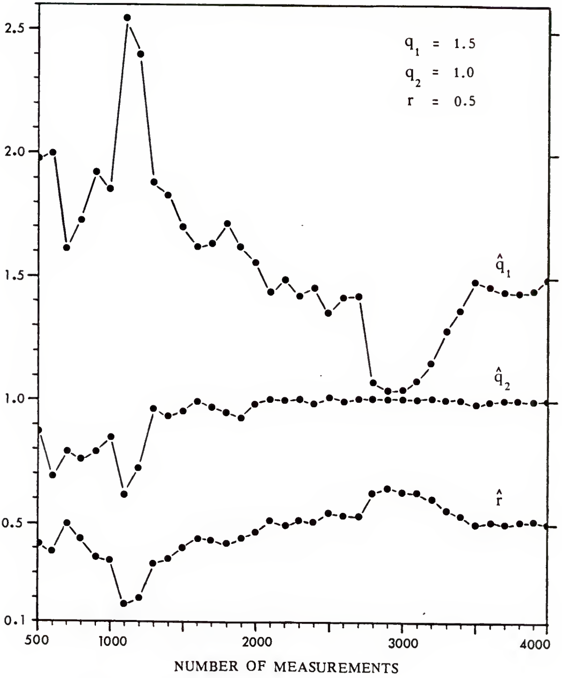

Figure 4.1: The Estimates of Noise Variances.

Based on Theorem 4.1 and Guassian properties of the noise sequences, the estimates of noise variances are expected to converge to

their actual value. Ten simulation runs with different noise sequences were made. The number of measurements used for each run was 4000. The simulation results are summarized in Table 4.1. The estimates of the noise variances q_1, q_2, and r for the first simulation run are depicted in Fig. 4.1 where the noise variance estimates are plotted for every new 100 data points. It is evident from Table 4.1 and Fig. 4.1 that the rate of convergence of $\hat{q}_1$ is rather slow in comparison with those of $\hat{q}_2$ and $\hat{r}$.

4.6 Concluding Remarks

A direct technique for estimating the unknown noise covariances which is referred to as the SPMC technique has been discussed and analyzed. By utilizing the minimal polynomial of the state matrix, we have described how stationary preprocessed measurements can be obtained. The unknown noise covariances can then be directly estimated from the sample autocovariances of the stationary preprocessed measurements without requiring the estimate of the state and the stationarity of the measurements.

It has been observed that stationary preprocessed measurements can be obtained by using any polynomial which is satisfied by the state matrix. But, the minimal polynomial gives the maximum number of independent equations which can be used to determine the unknown noise covariances. Under certain conditions on the fourth moments of the noise processes, it has been shown that the estimates of the noise covariances converge in quadratic mean to their actual value. Unique determinability of noise covariances has also been discussed. In particular, we have observed that the measurement noise covariance can

always be uniquely determined if the state matrix is nonsingular. In case that the state matrix is singular, one may or may not be able to uniquely determine the measurement noise covariance.

CHAPTER V
DETECTION OF ABRUPT CHANGES IN NOISE COVARIANCES

Unexpected changes often occur in dynamical systems. These changes may degrade the system performance. Such system changes are often termed "failures". There are several ways that failures can arise. Actuator failures and sensor failures are a few examples. To improve the performance of systems under failures, the failures must be promptly detected, isolated, and accommodated. We shall refer to a system which performs these three tasks as a failure detection system.

For the past decade, a wide variety of techniques have been proposed for designing failure detection systems. Extensive surveys on this subject can be found in Willsky (1976) and Isermann (1984). To design a failure detection system, the model of the corresponding dynamical system must be given. In some practical cases, uncertainties in system models arise. This leads to the issue of robust failure detection which has been investigated by Leininger (1981), Chow and Willsky (1984) and Lou et al. (1986).

Here, we are concerned with a special type of failures, namely, sensor failures. Sensor failures can occur in several forms. For example, sensors may become completely non-operational or unsatisfactorily operational. In the former case, sensors simply stop operating. While in the latter case the measurements provided by sensors contain unexpected disturbances. The degradation in sensor performance caused by these disturbances may be in the form of biases or

increased inaccuracies.

The degradation in sensor performance which we are concerned here is in the form of increased inaccuracies. As mentioned in Willsky (1976), this degradation can be modeled as abrupt increases in measurement noise covariances. For the degradation in the form of biases, readers may refer to Deckart et al. (1977) and Clark and Setzer (1980) for further discussions. The material presented in this chapter is a part of author's research reported in Bullock et al. (1987).

5.1 Problem Statement

Consider the discrete stochastic dynamical system described by

$$x_{k+1} = F_k x_k + G_k u_k + w_k \quad ; \quad x_o \sim N(\bar{x}_o, M_o) \ , \ w_k \sim N(0, Q_k) \qquad (5.1)$$

$$y_k = H_k x_k + v_k \quad ; \quad v_k \sim N(0, R_k) \text{ with } R_k > 0 \qquad (5.2)$$

where x_k, u_k , and y_k denote the state, the deterministic input, and the measurement, respectively. v_k and w_k denote respectively the system noise and measurement noise, assumed white. The x_o, $\{w_k\}$, and $\{v_k\}$ are assumed to mutually uncorrelated.

The estimate $\hat{x}_{k|k-1}$ of x_k given $\{y_o, y_1, \ldots, y_{k-1}\}$ can be obtained from the Kalman filter. With the above assumptions, $\hat{x}_{k|k-1}$ is the unbiased minimum variance estimate. Here we also assume that $(F_k, R_k^{-\frac{1}{2}} H_k)$ is uniformly detectable so that the one-step predictor error covariance M_k is uniformly bounded (Anderson and Moore, 1981). Let us now consider the case that sensors used to provide measurements can fail at some unknown time, say k_o . The effect of sensor failures is assumed to be in the form of increased inaccuracies and is modeled as increases in measurement noise covariances. With this assumption, the situation is that the filter has been designed for measurement noise

covariance R_k but yet the actual measurement noise covariance is $R_k^o > R_k$ for $k \geq k_o$. Given measurements y_k, we would like to determine if $k \geq k_o$.

We shall develop a residual-based detection scheme for the problem stated above. It should be noted that a standard voting scheme can also be employed for this problem. But the voting scheme requires at least three similar sensors in order to detect the occurrence of failures in one sensor. Because of its expensiveness in implementation, the voting scheme is thus unattractive.

5.2 The Effect of Sensor Failures on the Residual Sequence

Let us now investigate the effect of sensor failures on the residual sequence $\{\tilde{y}_{k|k-1}\}$. When the sensors have failed, R_k is still used in computing the Kalman filter gain K_k instead of the correct covariance, R_k^o . Consequently, M_k computed from the Kalman filter is not the actual one-step predictor error covariance. It should be noted that the estimate $\hat{x}_{k|k-1}$ obtained from the Kalman filter is no longer the minimum variance estimate of x_k for $k > k_o$. But $\hat{x}_{k|k-1}$ is still an unbiased estimate of x_k . It is important to observe that even though we may not have true divergence, i.e., the actual mean-squared error of the estimate becomes unbounded as time increases, in the situation described above, it is quite possible that the estimation errors may become intolerably large. This phenomenon which is known as apparent divergence (Fitzgerald, 1971) is equivalent in a practical sense to true divergence.

As discussed in Chapter II, the actual one-step predictor error convariance M_k^o satisfies the following difference equation.

$$M^o_{k+1} = \bar{F}_k M^o_k \bar{F}'_k + K_k R^o_k K'_k + Q_k \quad ; k \geq k_o \tag{5.3}$$

where $\bar{F}_k := F_k - K_k H_k$

Also
$$M_{k+1} = \bar{F}_k M_k \bar{F}'_k + K_k R_k K'_k + Q_k \tag{5.4}$$

Thus
$$M^o_{k+1} - M_{k+1} = \bar{F}(M^o_k - M_k)\bar{F}'_k + K_k(R^o_k - R_k)K'_k \quad ; k \geq k_o \tag{5.5}$$

Since $R^o_k > R_k$ and $M^o_{k_o} = M_{k_o}$, we then have that

$$M^o_k \geq M_k \quad ; k \geq k_o \tag{5.6}$$

It follows from (5.6) and $R^o_k > R_k$ that

$$\begin{aligned} \operatorname{COV}(\tilde{y}_{k|k-1}, \tilde{y}_{k|k-1}) &= H_k M_k H'_k + R_k \quad ; 0 \leq k < k_o \\ &> H_k M_k H'_k + R_k \quad ; k \geq k_o \end{aligned} \tag{5.7}$$

This suggests that determining if $k \geq k_o$ is equivalent to testing if the convariance of the residual sequence has increased. It should be pointed out that $\{\tilde{y}_{k|k-1}\}$ is white for $0 \leq k < k_o$, but it is no longer white for $k \geq k_o$. However, $\{\tilde{y}_{k|k-1}\}$ is a zero-mean Gaussian noise sequence for all $k \geq 0$ with covariance satisfying (5.7).

5.3 Hypothesis Testing for Sensor Failures

Since $(H_k M_k H'_k + R_k)$ is symmetric positive definite, there exists a nonsingular matrix T_k such that $T_k(H_k M_k H'_k + R_k) T'_k = I$.

Let us now define
$$z_k := T_k \tilde{y}_{k|k-1} \tag{5.8}$$

Then
$$z_k \sim N(0, V_k) \tag{5.9}$$

where
$$\begin{aligned} V_k &= I \quad ; 0 \leq k < k_o \\ &> I \quad ; k \geq k_o \end{aligned}$$

Since $\{\tilde{y}_{k|k-1}\}$ is white for $0 \le k < k_o$, z_o , z_1, ... , z_{k_o-1} are independent. Without sensor failures, one can therefore consider z_k as independent observations of a random vector **z** normally distributed with zero mean and identity covariance. The hypothesis test for sensor failures can be formulated as follows. Given n observations z_k of a random vector z normally distributed with zero mean and unknown covariance V. We perform the following multivariate hypothesis test.

$$\begin{aligned} H_o &: \text{ sensors have not failed} \quad V = I \\ H_1 &: \text{ sensors have failed} \quad V > I \end{aligned} \tag{5.10}$$

To obtain a test statistic for the above one-sided hypothesis test, we shall use Roy's union-intersection principle (Roy, 1953). The union-intersection principle is widely used in the area of simultaneous test procedures (Krishnaiah, 1979). However, it seems to be little known to engineers. The ideas of the union-intersection principle can be briefly stated as follows. Given a multivariate hypothesis H, we decompose H into a union or an intersection of several univariate hypotheses. For each univariate hypothesis, we develop a test statistic and obtain accordingly an acceptance region and a critical region. The acceptance region and the critical region of H are then obtained by appropriately intersecting or unioning all acceptance regions and critical regions of the univariate hypotheses. As mentioned in Roy (1957), if the test for each univariate hypothesis has certain optimum properties, then the test for H obtained by the above procedure will have some reasonably good properties.

Now observe that $V = I$ iff $\mathbf{a}' V \mathbf{a} = \mathbf{a}'\mathbf{a}$ for all $\mathbf{a}$ in $\mathbf{R}^p$, where p denotes the dimension of z. An equivalent hypothesis test to

(5.10) can then be given by

$$\begin{aligned} H_o &: \quad a' V a = a'a \quad \text{for all } a \text{ in } R^p \\ H_1 &: \quad a' V a > a'a \quad \text{for some } a \text{ in } R^p \end{aligned} \tag{5.11}$$

Let us next consider a subhypothesis test for a fixed non-zero a in R^p.

$$\begin{aligned} H_o^a &: \quad a' V a = a'a \\ H_1^a &: \quad a' V a > a'a \end{aligned} \tag{5.12}$$

Consider a test statistic $t(a) \quad := \quad \dfrac{a' A a}{a'a}$ (5.13)

where

$$A \quad := \sum_{i=0}^{n-1} z_i \, z'_i \tag{5.14}$$

It can be shown (see Appendix D) that t(a) has a chi-squared distribution with n degrees of freedom if $z \sim N(0,I)$. We may reject H_o^a if $t(a) \geq t_o$ where t_o is some positive number. It should be noted that the test based on this critical region is a uniformly most powerful test for given independent observations (see Appendix D).

Notice that

$$H_o = \bigcap_{a \neq o} H_o^a \quad \text{and} \quad H_1 = \bigcup_{a \neq o} H_1^a .$$

This suggests that we accept H_o if $t(a) < t_o$ for all non-zero a in R^p, or equivalently $\max_{a \neq o} t(a) < t_o$, and we reject H_o otherwise. Using the symmetric property of A, one can show that (Noble and Daniel, 1977):

$$\max_{a \neq o} t(a) \quad := \quad \max_{a \neq o} \frac{a' A a}{a'a} \quad = \quad \lambda_{max}(A) . \tag{5.15}$$

where λ_{max} (A) denotes the largest eigenvalue of A. Thus $\lambda_{max}(A)$ may be used as a test statistic for (5.10).

The random matrix A defined in (5.14) is well known to be distributed according to the Wishart distribution $W(I,n)$ under the null hypothesis H_o (Anderson, 1984). The distribution of the largest eigenvalue of A has a density function of a certain form depending on p and n. The exact form of the density function can be found in Krishnaiah and Chang (1971).

α	n \ p	2	3	4	5	6	7	8	9	10
0.05	2	8.59	10.74	12.68	14.49	16.21	17.88	19.49	21.06	22.61
	3	10.74	13.11	15.24	17.21	19.09	20.88	22.62	24.31	25.96
	4	12.68	15.24	17.52	19.63	21.62	23.53	25.37	27.15	28.90
	5	14.49	17.21	19.63	21.85	23.95	25.96	27.88	29.75	31.57
	6	16.21	19.09	21.62	23.95	26.14	28.23	30.24	32.18	34.08
	7	17.88	20.88	23.53	25.96	28.23	30.40	32.48	34.50	36.45
	8	19.49	22.62	25.37	27.88	30.24	32.48	34.63	36.70	38.72
	9	21.06	24.31	27.15	29.75	32.18	34.50	36.70	38.80	40.91
	10	22.61	25.96	28.90	31.57	34.08	36.45	38.72	40.91	43.04
	11	24.12	27.58	30.60	33.35	35.93	38.36	40.69	42.93	45.10
	12	25.61	29.17	32.27	35.09	37.73	40.22	42.60	44.90	47.12
	15	29.96	33.80	37.13	40.15	42.96	45.61	48.15	50.58	52.94
	20	36.94	41.18	44.84	48.14	51.21	54.10	56.86	59.50	62.05
0.01	2	12.16	14.57	16.73	18.73	20.64	22.47	24.23	25.95	27.63
	3	14.57	17.18	19.50	21.65	23.69	25.64	27.52	29.34	31.12
	4	16.73	19.50	21.96	24.24	26.38	28.43	30.41	32.32	34.18
	5	18.73	21.65	24.24	26.62	28.86	31.00	33.05	35.04	36.97
	6	20.64	23.69	26.38	28.86	31.19	33.40	35.53	37.59	39.59
	7	22.47	25.64	28.43	31.00	33.40	35.69	37.89	40.01	42.07
	8	24.23	27.52	30.41	33.05	35.53	37.89	40.15	42.33	44.45
	9	25.95	29.34	32.32	35.04	37.59	40.01	42.33	44.57	46.74
	10	27.63	31.12	34.18	36.97	39.59	42.07	44.45	46.74	48.95
	11	29.28	32.86	36.00	38.86	41.54	44.08	46.51	48.85	51.11
	12	30.89	34.56	37.78	40.71	43.45	46.04	48.52	50.91	53.22
	15	35.59	39.52	42.94	46.05	48.96	51.70	54.32	56.85	59.28
	20	43.08	47.37	51.10	54.49	57.63	60.60	63.43	66.14	68.77

Table 5.1: Upper Significance Points of $\lambda_{max}(A)$.

The approximate values of certain upper significance points of $\lambda_{max}(A)$ obtained from Pearson and Hartley (1972) are given in Table 5.1. Let λ_α be number such that $Pr\{\lambda_{max}(A) \geq \lambda_\alpha\} = \alpha$ under H_o. For example, if $p = 2$, $n = 10$, and $\alpha = 0.01$, then from Table 5.1 $\lambda_\alpha = 27.63$. We

then reject H_o and accept H_1 with significance level α if $\lambda_{max}(A)$ is greater or equal to λ_α .

So far, we have shown that the largest eigenvalue of random matrix $A = n.S$, where S is the sample covariance matrix based on n samples of normalized residual sequence, may be used as a test statistic in detecting sensor failures. However, a chi-squared test, which was suggested in Mehra and Peschon (1971) and applied by Willsky et al. (1975) to detect the occurrence of system abnormalities, can also be used to detect abrupt increases in measurement noise covariances. The chi-squared test is based on the trace of A. Under the null hypothesis H_o, tr(A) has a chi-squared distribution with n.p degrees of freedom. Let us now investigate the relationship between test statistics $\lambda_{max}(A)$ and tr(A).

Since A is symmetric, there exists an orthogonal matrix T such that

$$T' A T \quad = \quad D \qquad \text{where } D \text{ is diagonal.}$$

Now let $$\bar{a} \quad := \quad ([1\ 1\ .\ .\ 1]\ T'\ )' \tag{5.16}$$

Then $$\mathrm{tr}(A) \quad = \quad \bar{a}'\ A\ \bar{a}\ .$$

$$= \quad p.\frac{\bar{a}'\ A\ \bar{a}}{\bar{a}'\bar{a}} \qquad \text{since } \bar{a}'\bar{a} = p.$$

Consequently, $$\mathrm{tr}(A) \quad = p.t(\bar{a}) \tag{5.17}$$

It should be noted that $\bar{a}$ is itself a random vector since T is a random matrix depending on A. As far as the constant p is concerned, hypothesis tests based on $p.t(\bar{a})$ and $t(\bar{a})$ are equivalent. From (5.17), we can see that when we use tr(A) as a test statistic for (5.11), we are restricting our attention to subhypothesis test H_o^a against H_1^a with

$a = \bar{a}$. But in developing test statistic $\lambda_{max}(A)$ for (5.11), we have considered subhypothesis tests H_o^a against H_1^a for all nonzero a in R^p, which of course include the subhypothesis test $H_o^{\bar{a}}$ against $H_1^{\bar{a}}$. In this sense, test statistic $\lambda_{max}(A)$ is more general than test statistic tr(A) is.

5.4 A Scheme for Detecting Sensor Failures

Let us now develop a scheme for detecting sensor failures. Given $\{\tilde{y}_{k|k-1}\}$, let n now denote the value of the window length used in the detection scheme, and let k denote time index when the detection scheme is performed with $k \geq n-1$.

Define
$$A_k := \sum_{i=k-n+1}^{k} z_i z_i' \tag{5.18}$$
$$\lambda_k := \lambda_{max}(A_k) \tag{5.19}$$

where z_i is defined in (5.8).

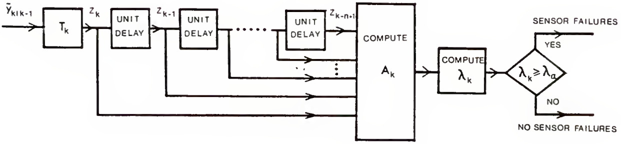

Figure 5.1: Sensor Failure Detection.

Let α be the significance level used in the hypothesis test, and λ_α be the threshold corresponding to α. The detection scheme is that if $\lambda_k < \lambda_\alpha$, we decide that the sensors have not failed up to time k, or else we decide that the sensors have failed at some particular time before or at time k. A block diagram illustrating the detection scheme is depicted in Fig. 5.1.

After sensor failures are detected, one needs to adjust the filter gain K_k to accommodate the failures. Here, we propose a simple scheme for the adjustment. The idea is to obtain the estimate of $\mathrm{COV}(\tilde{y}_{k|k-1},\tilde{y}_{k|k-1})$ and then adjust the filter gain K_k accordingly. Based on the assumption that $\mathrm{COV}(\tilde{y}_{k|k-1},\tilde{y}_{k|k-1})$ is constant over N consecutive samples, the sample covariance of $\tilde{y}_{k|k-1}$, denoted by S_k, is computed and used as an estimate of $\mathrm{COV}(\tilde{y}_{k|k-1},\tilde{y}_{k|k-1})$. Assume further that $(M_k^o - M_k)$ is not considerably large when the failures are detected. The filter gain K_k may then be given by

$$K_k \quad = \quad F_k\, M_k\, H_k'\, S_k^{-1} \tag{5.20}$$

where
$$S_k \quad := \quad \frac{1}{N} \sum_{i=k-N+1}^{k} \tilde{y}_{k|k-1}\, \tilde{y}_{k|k-1}' \tag{5.21}$$

It should be noted that adjusting K_k according to (5.20) is equivalent to adjusting K_k according to (5.22), where $\hat{R}_k^o$ may be interpreted as an estimate of the measurement noise covariance.

$$K_k \quad = \quad F_k\, M_k\, H_k'\, (H_k M_k H_k' + \hat{R}_k^o)^{-1} \tag{5.22}$$

where
$$\hat{R}_k^o \quad := \quad S_k - H_k\, M_k\, H_k' \tag{5.23}$$

It is important to observe that when K_k is computed from (5.20), the filter error covariance P_k , which is approximated by (5.24), may not be positive semi-definite.

$$P_k = M_k - M_k H_k' S_k^{-1} H_k M_k \tag{5.24}$$

A sufficient condition for P_k being positive semi-definite is that $(S_k - H_k M_k H_k') \geq 0$. This condition does not always hold since S_k is only the sample covariance of $\tilde{y}_{k|k-1}$. But from (5.7), we know that

$$\mathrm{COV}(\tilde{y}_{k|k-1}, \tilde{y}_{k|k-1}) > H_k M_k H_k' + R_k \qquad \text{for } k \geq k_o .$$

The above information about the covariance of $\tilde{y}_{k|k-1}$ should also be utilized in adjusting the filter gain K_k. Thus, we may simply replace S_k by $H_k M_k H_k' + R_k$ when $(S_k - H_k M_k H_k' - R_k)$ is not positive semi-definite. By doing this, P_k is guaranteed to be positive semi-definite at the same time. The sensor failure detection system described above has been reported by Bullock et al. (1987) based on their Monte-Carlo study to work satisfactorily.

It should be noted that sensor failures may also be accommodated by simply turning off the failed sensors. But this may make the system become unobservable. Once the system is unobservable, the state error covariance may become unbounded as time increases. In particular, if the state transition matrix of the system is not exponentially stable, the state error covariance can be intolerably large within a short period of time after the system becomes unobservable.

5.5 Concluding Remarks

Abrupt increases in measurement noise covariances due to sensor failures in the form of increased inaccuracies have been considered. It has been shown that this type of sensor failures affects the residual sequence by an unknown increment in its covariances. Consequently, the sensor failures can be detected by simply testing whether or not the

covariance of the residual sequence has increased. Utilizing Roy's union-intersection principle, we have shown that the test statistic which can be used in the detection is the largest eigenvalue of random matrix $A = n.S$, where S denotes the sample covariance matrix based on n samples of the normalized residual sequence.

The sensor failure detection developed in this chapter can be directly applied to formulate an adaptive state estimation scheme which is robust against sensor failures. In state estimation, one can improve the robustness against sensor failures by simply using several different types of sensors. Since the sensors are implemented with dissimilar instruments, they are unlikely to fail at the same time. For instance, an active radar seeker and an imaging infrared seeker are usually used in target tracking. However, failure in one sensor can deteriorate the overall state estimate. For this reason, one might want to use sensor failure detection to identify the failed sensor and accommodate the failure accordingly.

CHAPTER VI
CONCLUSIONS

This dissertation has analyzed Kalman filtering under uncertainty in noise covariances. Particular emphasis has been given to discrete-time Kalman filtering. The dissertation consists of three main parts.

In the first part, we considered situations in which the Kalman filter is designed by using incorrect noise covariances. Behavior of the Kalman filter under incorrect noise covariances was analyzed. Particularly, we were interested in the characteristic of the actual performance of the filter. The filter performance was quantified by the actual error covariance of the state estimate. Through this quantity, the characteristic of the filter was studied. In particular, two important properties of the actual error covariance, convergence and divergence, were investigated. The contribution of the results developed in this part is that they help one to understand and be able to predict certain behavior of the Kalman filter when inexact values of noise covariances are used. This is, of course, important since the exact values of the noise covariances are hardly known in most practical cases.

In the second part, situations in which the noise convariance are unknown were considered. A direct technique for estimating the unknown noise covariances which was referred to as the SPMC technique was discussed and analyzed. The technique is direct in the sense that the unknown noise covariance can be estimated without requiring the estimate

of the state and the stationarity of the measurements. A sufficient condition for which the estimates of the noise covariances converge in quadratic mean to their actual value was established. In addition, discussions on unique determinability of the unknown noise covariances were given.

In the third part, we considered cases in which the noise covariances are subject to abrupt changes. Particularly, we were interested in abrupt increases in measurement noise covariances which were used to model the effect of sensor failures in the form of increased inaccuracies. It was shown that this type of sensors failures affected the residual sequence by an unknown increment in its covariance. The sensor failures can therefore be detected by simply testing whether or not the covariance of the residual sequence has increased. It was shown that the test statistic which could be used in the detection was the largest eigenvalue of random matrix $A = n.S$ where S denotes the sample covariance matrix based on n samples of the normalized residual sequence. A scheme for detecting abrupt increases in measurement noise covariances was then developed in this part.

APPENDIX A
STABILITY OF LINEAR SLOWLY TIME-VARYING SYSTEMS

In this appendix, we briefly discuss stability of linear slowly time-varying systems. First, let us consider the continuous-time system described by

$$\dot{x} = A(t)\,x \tag{A.1}$$

It is well known that stability of the above time-varying system, in general, cannot be concluded from stability of the corresponding frozen-time systems. To see this, we consider the following two examples.

<u>Example A.1</u> (Vidyasagar, 1978): Let A(t) be given by

$$A(t) = \begin{bmatrix} -1 + a.\cos^2 t & 1 - a.\sin t.\cos t \\ -1 - a.\sin t.\cos t & -1 + a.\sin^2 t \end{bmatrix}$$

Then, the corresponding state transition matrix $\Phi(t,0)$ is given by

$$\Phi(t,0) = \begin{bmatrix} e^{(a-1)t}\cos t & e^{-t}\sin t \\ -e^{(a-1)t}\sin t & e^{-t}\cos t \end{bmatrix}$$

Observe that the eigenvalues of A(t) are independent of t and satisfy the following characteristic equation.

$$s^2 + (2 - a)s + (2 - a) = 0$$

Now, if $a < 2$, then A(t) has all its eigenvalues with negative real parts for all t. It is however clear from the above expression of

$\Phi(t,0)$ that the system (A.1) is unstable for all $a > 1$. This shows that stability of the system (A.1), in general, cannot be predicted from stability of the frozen-time systems, i.e., from the eigenvalues of $A(t)$.

Example A.2 (Skoog and Lau, 1972): Let $A(t)$ be given by

$$A(t) \quad = \quad \begin{bmatrix} -1 + a.\sin t.\cos t & 1 + a.\cos^2 t \\ -1 - a.\sin^2 t & -1 - a.\sin t.\cos t \end{bmatrix}$$

The corresponding state transition matrix $\Phi(t,0)$ can be expressed as

$$\Phi(t,0) \quad = \quad \begin{bmatrix} e^{-t}\cos t & e^{-t}(\sin t + a.t.\cos t) \\ -e^{-t}\sin t & e^{-t}(\cos t - a.t.\sin t) \end{bmatrix}$$

It can be shown that the eigenvalues of $A(t)$ are independent of t and satisfy the following characteristic equation.

$$s^2 + 2s + (2 + a) \quad = \quad 0$$

For, $a < -2$, $A(t)$ has an eigenvalue with a positive real part. But yet, the system (A.1) is exponentially stable for all a. Thus, this example shows that instability of the system (A.1), in general, can neither be predicted from the eigenvalues of $A(t)$.

In case that the variation of $A(t)$ is sufficiently slow, stability of the system (A.1) can however be predicted from the eigenvalues of $A(t)$. This fact was first pointed out by Rosenbrock (1963). The following theorem which has been used in Chapter 3 is due to Desoer (1969).

Theorem A.1: Let $A(t)$ be differentiable for all $t \geq 0$. Suppose that

(i) $A(t)$ is uniformly bounded, and

(ii) $A(t)$ has all its eigenvalues with real parts less than or

equal to $-\sigma$ $\forall\, t \geq 0$, for some positive number σ.

Then, there exists a positive number ϵ such that if $\sup_{t\geq 0} \|\dot{A}(t)\| \leq \epsilon$, the system (A.1) is exponentially stable.

For discrete-time system

$$x_{k+1} = A_k\, x_k \tag{A.2}$$

Desoer (1970) has proven the following theorem which is the discrete-time counterpart of Theorem A.1.

Theorem A.2: Suppose that

(i) A_k is uniformly bounded, and

(ii) A_k has all its eigenvalues inside or on the circle, centered at the origin, of radius ρ $\forall\, k \geq 0$, for some $\rho \in (0,1)$.

Then, there exists a positive number ϵ such that if $\sup_{k\geq 0} \|A_{k+1} - A_k\| \leq \epsilon$, the system (A.2) is exponentially stable.

APPENDIX B
CONVERGENCE PROPERTY OF SOLUTIONS OF THE RICCATI DIFFERENCE EQUATION

In this appendix, we shall prove the convergence property of solutions of the **RDE** stated in the proof of Corollary 2.3. Let us now consider the **RDE** given by (2.3) with constant F, H, Q, and R. The following lemma demonstrates how solutions of the **RDE** with different initial values are related.

<u>Lemma B</u>: Let $M_{k,1}$ and $M_{k,2}$ be solutions of the **RDE** with different initial values. Then,

$$\Delta_{k+1} = \bar{F}_{k,1}\, \Delta_k\, \bar{F}'_{k,2} \tag{B.1}$$

where $\Delta_k := M_{k,1} - M_{k,2}$ and $\bar{F}_{k,j} := F - K_{k,j}H$

with $K_{k,j} := FM_{k,j}H'(HM_{k,j}H' + R)^{-1}$, for $j = 1,2$.

<u>Proof</u>: Observe that

$$M_{k+1,1} = \bar{F}_{k,1}M_{k,1}F' + Q \quad \text{and} \quad M_{k+1,2} = FM_{k,2}\bar{F}'_{k,2} + Q .$$

Thus, $\Delta_{k+1} = \bar{F}_{k,1}M_{k,1}F' - FM_{k,2}\bar{F}'_{k,2}$

$$= \bar{F}_{k,1}\Delta_k\, \bar{F}'_{k,2} + \bar{F}_{k,1}M_{k,1}H'K'_{k,2} - K_{k,1}HM_{k,2}\bar{F}'_{k,2}$$

It can be shown that the sum of the last two terms on the RHS of the above equation is equal to zero. Equation (B.1) therefore holds. ∎

Let us now prove the following convergence property of solutions of the **RDE**.

<u>Theorem B</u>: Let (F,H) be detectable and $(F,Q^{1/2})$ have no unreachable mode on the unit circle. Suppose that the solution of the **RDE** with initial value $\Pi \geq 0$ converges to the stabilizing solution M of the **ARE**. Then, the solution of the **RDE** with initial value $\alpha.\Pi$ for any positive number α also converges to M.

<u>Proof</u>: First notice that the existence of the stabilizing solution M is guaranteed by the above assumptions on F,H, and Q. Let $M_{k,1}$ and $M_{k,2}$ be solutions of the **RDE** with initial values Π and $\alpha.\Pi$, respectively. From Lemma B, we have that

$$\Delta_{k+1} = \bar{F}_{k,1}\, \Delta_k\, \bar{F}'_{k,2}$$

Hence,

$$\Delta_k = (1-\alpha)\Psi_1(k,0)\Pi\ \Psi_2'(k,0) \tag{B.2}$$

where $\Psi_j(k,i)$ denotes the state transition matrix associated with $\bar{F}_{k,j}$ for $j = 1,2$.

But $M_{k,2} = \alpha.\ \Psi_2(k,0)\Pi\ \Psi_2'(k,0)$ + nonnegative definite terms.

Hence

$$M_{k,2} \geq \alpha.\ \Psi_2(k,0)\Pi\ \Psi_2'(k,0) \tag{B.3}$$

Because of the detectability of (F,H), $M_{k,2}$ is uniformly bounded. It then follows from (B.3) that $\Pi\ \Psi_2'(k,0)$ is uniformly bounded. But from Remark 2.4, we have that $\bar{F}_{k,1}$ is exponentially stable. It is now clear from (B.2) that $\lim_{k\to\infty} \Delta_k = 0$. Therefore, $\lim_{k\to\infty} M_{k,2} = M$. ∎

<u>Remark B</u>: As stated in Remark 2.5, the solution of the **RDE** with initial value $\Pi \geq M$, where M denotes the stabilizing solution of the **ARE**, converges to M. Based on this fact and the above theorem, one can conclude that the solution of the **RDE** with initial value Π such that $\alpha.\Pi \geq M$ for some positive number α also converges to M.

APPENDIX C
CONVERGENCE PROPERTY OF
SOLUTIONS OF THE RICCATI DIFFERENTIAL EQUATION

Consider the **RDE** given by (3.5) with constant F,H,Q, and R. In this appendix, we shall prove two convergence properties of solutions of the **RDE**, which have been stated in Section 3.2. First, let us prove the following lemma which demonstrates how solutions of the **RDE** with different initial values are related.

<u>Lemma C</u>: Let $P_1(t)$ and $P_2(t)$ be solutions of the **RDE** with different initial values. Then,

$$\dot{\Delta}(t) = \overline{F}_1(t)\Delta(t) + \Delta(t)\overline{F}_2'(t) \tag{C.1}$$

where $\Delta(t) := P_1(t) - P_2(t)$

and $\overline{F}_i(t) := F - K_i(t)H$ with $K_i := P_i(t)H'R^{-1}$, for $i = 1,2$.

<u>Proof</u>:

$$\begin{aligned}\dot{\Delta}(t) &= \overline{F}_1(t)P_1(t) + P_1(t)\overline{F}_1'(t) + K_1(t)RK_1'(t) + Q \\ &\quad - \overline{F}_2(t)P_2(t) - P_2(t)\overline{F}_1'(t) - K_2(t)RK_2'(t) - Q \\ &= \overline{F}_1(t)\Delta(t) + \Delta(t)\overline{F}_2'(t) \\ &\quad + P_1(t)\overline{F}_1'(t) - P_1(t)\overline{F}_2'(t) + K_1(t)RK_1'(t) \\ &\quad + \overline{F}_1(t)P_2(t) - \overline{F}_2(t)P_2(t) - K_2(t)RK_2'(t)\end{aligned}$$

It can be shown that the sum of the last six terms on RHS of the above equation is equal to zero. Thus, $\Delta(t)$ satisfies (C.1). ∎

The convergence properties of solutions of the **RDE** are given by the following two theorems.

Theorem C.1: Let (F,H) be detectable and $(F,Q^{1/2})$ have no uncontrollable mode on the imaginary axis. If $P(0) > 0$, then $P(t)$ converges to the unique stabilizing solution P of the **ARE**.

Proof: Under the above assumptions on F,H, and Q, the stabilizing solution P of the **ARE** exists and is unique. Using Lemma C, we have that

$$\dot{\Delta}(t) = \bar{F}(t)\Delta(t) + \Delta(t)\bar{F}' \tag{C.2}$$

where $\Delta(t) := P(t) - P$

Hence,
$$\Delta(t) = \Psi(t,0)(P(0) - P)\, e^{\bar{F}'t} \tag{C.3}$$

where $\Psi(t,s)$ is the state transition matrix associated with $\bar{F}(t)$.

But $$P(t) = \Psi(t,0)P(0)\Psi'(t,0) + \int_0^t \Psi(t,s)[K(s)RK'(s) + Q]\Psi'(t,s)ds$$

Thus $$P(t) \geq \Psi(t,0)P(0)\Psi'(t,0) \tag{C.4}$$

Since (F,H) is detectable, $P(t)$ is uniformly bounded. Because of uniform boundedness of $P(t)$ and positive definiteness of $P(0)$, inequality (C.4) implies that $\Psi(t,0)$ is uniformly bounded. Using uniform boundedness of $\Psi(t,0)$ and exponential stability of $\bar{F}$, we obtain from (C.3) that $\Delta(t)$ converges to a zero matrix. Consequently, $\lim_{t\to\infty} P(t) = P$. ■

Theorem C.2: Let (F,H) be detectable and $(F,Q^{1/2})$ have no uncontrollable mode on the imaginary axis. Suppose that the solution of the **RDE** with initial value $\Pi \geq 0$ converges to the unique stabilizing solution P of the **ARE**. Then the solution of the **RDE** with initial value $\alpha.\Pi$ for any positive number α also converges to P.

Proof: Let $P_1(t)$ and $P_2(t)$ be solutions of the **RDE** with initial values Π and $\alpha.\Pi$, respectively. By Lemma C,

$$\dot{\Delta}(t) \quad = \quad \overline{F}_1(t)\Delta(t) \quad + \quad \Delta(t)\overline{F}_2'(t)$$

Thus $$\Delta(t) \quad = \quad (1 - \alpha).\Psi_1(t,0)\Pi \; \Psi_2'(t,0) \tag{C.5}$$

where $\Psi_i(t,s)$ denotes the state transition matrix associated with $\overline{F}_i$ for $i = 1, 2$.

Using the same arguments given in the proof of Theorem C.1, we have that $\Psi_2(t,0) \; \Pi \; \Psi_2'(t,0)$ is uniformly bounded. This in turn implies that $\Pi \; \Psi_2'(t,0)$ is uniformly bounded. From Remark 3.5, we have that $\Psi_1(t,0)$ is exponentially stable. It is then clear from (C.5) that $\lim_{t\to\infty} \Delta(t) = 0$. Thus, $P_2(t)$also converges to P. ∎

APPENDIX D
DISTRIBUTION AND PROPERTY OF TEST STATISTIC t(a)

Let $z_1, z_2, \ldots, z_n$ be n independent observations of a p-dimensional random vector $z \sim N(0,I)$. Let us first prove that t(a) defined below has a chi-squared distribution with n degrees of freedom.

$$t(a) \quad := \quad \frac{a' \, A \, a}{a'a}$$

where $\quad A \quad := \sum_{i=1}^{n} z_i \, z_i'$ $\quad$ and a is a nonzero vector in R^p.

Let us state the following two theorems (Lancaster, 1969) needed for completing the proof.

<u>Theorem D.1</u>: Let x and y be independent random variables distributed according to chi-squared distributions with k and m degrees of freedom, respectively. Then x + y has a chi-squared distribution with k + m degrees of freedom.

<u>Theorem D.2</u>: Let p-dimensional random vector $x \sim N(0,I)$, B be a pxp symmetric matrix, and $m \leq p$. Then $x' \, B \, x$ has a chi-squared distribution with m degrees of freedom iff B is idempotent of rank m.

Let us now prove that t(a) has a chi-squared distribution with n degrees of freedom. Based on Theorem D.1, it is sufficient to show that $a'zz'a \, / \, a'a$ has a chi-squared distribution with one degree of freedom.

From $\dfrac{a'zz'a}{a'a} = \dfrac{z'aa'z}{a'a} = z' \, B \, z$ where $B := \dfrac{1}{a'a} \, aa'$.

It can be observed that B is idempotent of rank 1. By Theorem D.2 , $a'zz'a / a'a$ has a chi-squared distribution with one degree of freedom. Consequently, $t(\mathbf{a}) := \sum_{i=1}^{n} a'z_i z_i' a / a'a$ has a chi-squared distribution with n degrees of freedom.

Given $z_1, z_2, \ldots, z_n$ independent observations of random vector **z** normally distributed with mean zero and unknown covariance V. It is however known a priori that either V = I or V > I holds. Next, we shall show that for a subhypothesis test with fixed non-zero vector **a** :

$$H_o^a \quad : \quad a' \, V \, a \quad = \quad a'a$$

$$H_1^a \quad : \quad a' \, V \, a \quad > \quad a'a$$

the test based on a critical region $\{ t(\mathbf{a}) \geq t_\alpha \}$ is a uniformly most powerful test with significance level α , where t(a) is defined above and t_α is a number such that $\Pr\{ t(\mathbf{a}) \geq t_\alpha \} = \alpha$. Roughly, this means that the probability of rejecting the null hypothesis when it is false of this test is greater than or equal to that of any test with significance level less then or equal to α . The exact definition of a uniformly most powerful test with significance level α can be found in Mood et al. (1974) or Lehmann (1986). To prove the above statement, we need the following result (Mood et al., 1974).

Theorem D.3: Let $x_1, \ldots, x_n$ be independent observations of random variable **x** whose density function is given by

$$p_{\mathbf{x}}(x;\theta) \quad = \quad a(\theta).b(x).\exp[c(\theta).d(x)] \quad \text{for some parameter } \theta.$$

Let $t(x_1, \ldots, x_n) := \sum_{i=1}^{n} d(x_i)$. If $c(\theta)$ is a monotone, increasing function in θ and there exists t^* such that

$$\Pr\{ t(x_1, \ldots, x_n) \geq t^* \mid \theta = \theta_o \} \quad = \quad \alpha ,$$

then the test based on critical region $\{ t(x_1,\ldots,x_n) \geq t^* \}$ is a uniformly most powerful test with significance level α of

$$H_o: \theta = \theta_o \quad \text{against} \quad H_1: \theta > \theta_o .$$

Now let

$$\mathbf{x} := \mathbf{a'zz'a} \,/\, \mathbf{a'a}$$

$$\theta := \mathbf{a'} \mathrm{V} \mathbf{a}$$

$$\theta_o := \mathbf{a'a} \quad .$$

Then

$$\mathbf{x} = \frac{\theta}{\theta_o} \mathbf{w} \qquad \text{where} \quad \mathbf{w} := \frac{\mathbf{a'zz'a}}{\mathbf{a'} \mathrm{V} \mathbf{a}} .$$

Since V is symmetric positive definite, there exists an invertible matrix T such that T V T′ = I. Let $\mathbf{y} := \mathrm{T}\mathbf{z}$, then $\mathbf{y} \sim N(0,I)$. Since T is invertible, there is a non-zero **b** such that $\mathbf{a} = \mathrm{T'}\mathbf{b}$.

$$\text{Hence,} \quad \frac{\mathbf{a'zz'a}}{\mathbf{a'} \mathrm{V} \mathbf{a}} = \frac{\mathbf{b'Tzz'T'b}}{\mathbf{b'TVT'b}} = \frac{\mathbf{b'yy'b}}{\mathbf{b'b}} .$$

But $\mathbf{b'yy'b} / \mathbf{b'b}$ has a chi-squared distribution with one degree of freedom. Thus **w** has a chi-squared distribution with one degree of freedom, and its density function is given by

$$p_{\mathbf{w}}(w) = \begin{cases} \exp[-w/2]/(2\pi\, w)^{1/2} & , w \geq 0 \\ 0 & , w < 0 \end{cases}$$

$$\text{But} \quad p_{\mathbf{x}}(x) = p_{\mathbf{w}}(f^{-1}(x)) . \left| \frac{df^{-1}}{dx}(x) \right| \quad \text{where} \quad f(\mathbf{w}) := \frac{\theta}{\theta_o} \mathbf{w} .$$

$$\text{Thus} \quad p_{\mathbf{x}}(x) = a(\theta).b(x).\exp[c(\theta).d(x)]$$

where

$$a(\theta) := 1/(2\pi\theta \,/\, \theta_o)^{1/2}, \qquad b(x) := x^{-1/2},$$

$$c(\theta) := -(\theta_o / \theta)/\, 2 \quad , \quad \text{and} \quad d(x) := x.$$

It is clear that $c(\theta)$ is a monotone, increasing function in θ,

and $t(\mathbf{a}) = \sum_{i=1}^{n} d(x_i)$. By Theorem D.3, the test with critical region $\{ t(a) \geq t_\alpha \}$ is a uniformly most powerful test with significance level α of

$$H_o^a : \mathbf{a'} \mathrm{V} \mathbf{a} = \mathbf{a'a} \quad \text{against} \quad H_1^a : \mathbf{a'} \mathrm{V} \mathbf{a} > \mathbf{a'a} .$$

REFERENCES

Aasnaes, H.B. and T. Kailath (1974). "Initial-condition robustness of linear least squares filtering algorithms," IEEE Trans. Automat. Contr., vol. AC-19, pp. 393-397.

Alspach, D.L (1972). "Comments on: On the identification of variances and adaptive Kalman filtering," IEEE Trans. Automat. Contr., vol. AC-17, pp. 843-845.

Anderson, B.D.O. (1971a). "Stability properties of Kalman-Bucy filters," J. Franklin Inst., vol 291, pp. 137-144.

Anderson, B.O.D. and J.B. Moore (1979). Optimal Filtering, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

Anderson, B.O.D. and J.B. Moore (1981). "Detectability and stabilizability of time varying discrete time linear systems," SIAM J. Contr. Optimiz., vol. 19, pp. 20-32.

Anderson, T.W. (1971). The Statistical Analysis of Time Series, John Wiley & Sons, New York.

Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, New York.

Aoki, M. (1987). State Space Modeling of Time Series, Springer-Verlag, Heidelberg, Germany.

Belanger, P.R. (1974). "Estimation of noise covariance matrices for a linear time-varying stochastic process," Automatica, vol. 10, pp. 267-275.

Box, G.E.P. and G.M. Jenkins (1976). Time Series Analysis: Forecasting and Control, Holden-Day, Oakland, California.

Brewer, J.W. (1978). "Kronecker products and matrix calculus in system theory," IEEE Trans. Circuits Syst., vol. CAS-25, pp. 772-781.

Brown, R.J. and A.P. Sage (1971a). "Error analysis of modeling and bias errors in continuous time state estimation," Automatica, vol. 7, pp. 577-590.

Brown, R.J. and A.P. Sage (1971b). "Analysis of modeling and bias errors in discrete-time state estimation," IEEE Trans. Aerosp. Electron. Syst., vol. AES-7, pp. 340-354.

Bullock, T.E., S. Sangsuk-Iam, and T. Yoon (1987). "Sensor fusion study: Modification to experimental study of recent generalization of optical matched filtering," Final Report, Air Force Armament Laboratory, Elgin Air Force Base, Florida.

Burg, J.P., D.G. Luenberger, and D.L. Wenger (1982). "Estimation of structured covariance matrices," Proc. IEEE, vol. 70, pp. 963-974.

Caines, P.E. and D.Q. Mayne (1970). "On the discrete time matrix Riccati equation of optimal control," Int. J. Contr., vol. 12, pp. 785-794.

Caines, P.E. and D.Q. Mayne (1971). "On the discrete time matrix Riccati equation of optimal control--a correction," Int. J. Contr., vol. 14, pp. 205-207.

Callier, F.M. and J.L. Willems (1981). "Criterion for convergence of the solution of the Riccati differential equation," IEEE Trans. Automat. Contr., vol. AC-26, pp. 1232-1242.

Carew, B. and P.R. Belanger (1973). "Identification of optimum filter steady-state gain for systems with unknown noise covariances," IEEE Trans. Automat. Contr., vol. AC-18, pp. 582-587.

Chan, S.W., G.C. Goodwin, and K.S. Sin (1984). "Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems," IEEE Trans. Automat. Contr., vol. AC-29, pp. 110-118

Chow, E.Y. and A.S. Willsky (1984). "Analytical redundancy and the design of robust failure detection systems," IEEE Trans. Automat. Contr., vol. AC-29, pp. 603-614.

Chung, K.L. (1974). A Course in Probability Theory, Academic Press, New York.

Clark, R.N. and W. Setzer (1980). "Sensor fault detection in a system with random disturbances," IEEE Trans. Aerosp. Electron. Syst., vol. AES-16, pp. 468-473.

Deckert, J.C., D.N. Mukund, J.J. Deyst, and A.S. Willsky (1977). "F-8 DFBW sensor failure identification using analytic redundancy," IEEE Trans. Automat. Contr., vol. AC-22, pp. 795-803.

Desoer, C.A. (1969). "Slowly varying system $\dot{x} = A(t)x$," IEEE Trans. Automat. Contr., vol. AC-14, pp. 780-781.

Desoer, C.A. (1970). "Slowly varying discrete system $x_{i+1} = A_i x_i$," Electronics Letters, vol. 6, pp. 339-340.

De Souza, C.E., M.R. Gevers, and G.C. Goodwin (1986). "Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices," IEEE Trans. Automat. Contr., vol. AC-31, pp. 831-838.

Fagin, S.L. (1964). "Recursive linear regression theory, optimal filtering theory, and error analyses of optimal systems," IEEE Int. Conv. Rec., pp. 216-240.

Fitzgerald, R.J. (1971). "Divergence of the Kalman filter," IEEE Trans. Automat. Contr., vol. AC-16, pp. 736-747.

Friedland, B. (1982). "Estimating noise variances by using multiple observers," IEEE Trans. Aerosp. Electron. Syst., Vol. AES-18, pp. 442-448.

Fujita, S. and T. Fukao (1970). "Error analysis of optimal filtering algorithm for colored noise," IEEE Trans. Automat. Contr., vol. AC-15, pp. 452-455.

Fuller, W.A. (1976). Introduction to Statistical Time Series, John Wiley & Sons, New York.

Griffin, R.E. and A.P. Sage (1968). "Large and small scale sensitivity analysis of optimum estimation algorithms," IEEE Trans. Automat. Contr., vol. AC-13, pp. 320-329.

Griffin, R.E. and A.P. Sage (1969). "Sensitivity analysis of discrete filtering and smoothing algorithms," AIAA J., vol. 7, pp. 1890-1897.

Heffes, H. (1966). "The effect of erroneous models on the Kalman filter response," IEEE Trans. Automat. Contr., vol. AC-11, pp. 541-543.

Herstein, I.N. (1975). Topics in Algebra, John Wiley & Sons, New York.

Hoeffding, W. and H. Robbins (1948). "The central limit theorem for dependent random variables," Duke Math. J., vol. 15, pp. 773-780.

Huddle, J.R. and D.A. Wismer (1968). "Degradation of linear filter performance due to modeling error," IEEE Trans. Automat. Contr., vol. AC-13, pp. 421-423.

Iglehart, S.C. and C.T. Leondes (1974). "Estimation of a dispersion parameter in discrete Kalman filtering," IEEE Trans. Automat. Contr., vol. AC-19, pp. 262-263.

Isermann, R. (1984). "Process fault detection based on modeling and estimation methods - A survey," Automatica, vol. 20, pp. 387-404.

Jazwinski, A.H. (1970). Stochastic Processes and Filtering Theory, Academic Press, New York.

Kailath, T. (1968). "An innovations approach to least-squares estimation: Part I, Linear filtering in additive white noise," IEEE Trans. Automat. Contr., vol. AC-13, pp. 646-655.

Kalman, R.E. (1960). "A new approach to linear filtering and prediction problems," ASME J. Basic Eng., vol. 82, pp. 35-45.

Kalman, R.E. (1963). "New methods in Wiener filtering theory," Proc. Symp. Eng. Appl. Random Function Theory and Probability, Boydanoff, J.L. and F. Kozin Eds., John Wiley, New York.

Kalman, R.E. and R.S. Bucy (1961). "New results in filtering and prediction theory," ASME J. Basic Eng., vol. 83, pp. 95-108.

Kato, T. (1976). Perturbation Theory for Linear Operators, Springer-Verlag, New York.

Koussoulas, K.T. and C.T. Leondes (1986). "On the sensitivity of a discrete-time Kalman filter to plant-dynamics modelling errors," Int. J. Syst. Sci., vol. 17, pp. 937-941.

Krishnaiah, P.R. (1979). "Some developments on simultaneous test procedures," Chapter 4 in Developments in Statistics, vol. 2, Academic Press, New York.

Krishnaiah, P.R. and T.C. Chang (1971). "On the exact distribution of the extreme roots of the Wishart and MANOVA Matrices," J. Multivariate Analysis, vol. 1, pp. 108-117.

Kucera, V. (1972a). "A contribution to matrix quadratic equations," IEEE Trans. Automat. Contr., vol. AC-17, pp. 344-357.

Kucera, V. (1972b). "The discrete Riccati equation of optimal control," Kybernetika, vol. 8, pp. 430-447.

Kucera, V. (1973). "A review of the matrix Riccati equation," Kybernetika, vol. 9, pp. 42-61.

Kwakernaak, H. and R. Sivan (1972). Linear Optimal Control Systems, Wiley-Interscience, New York.

Lainiotis, D.G. and F.L. Sims (1970). "Sensitivity analysis of discrete Kalman filters," Int. J. Contr., vol. 12, pp. 657-669.

Lancaster, H.O. (1969). The Chi-Squared Distribution, John Wiley & Sons, New York.

Lee, D.P., S.E. Cohn, A. Dalcher, and M. Ghil (1985). "An efficient algorithm for estimating noise covariances in distributed systems," IEEE Trans. Automat. Contr., vol. AC-30, pp. 1057-1064.

Lee, T.T. (1980). "Direct approach to identify the noise covariances of Kalman filtering," IEEE Trans. Automat. Contr., vol. AC-25, pp. 841-842.

Lehmann, E.L. (1986). Testing Statistical Hypotheses, John Wiley & Sons, New York.

Leininger, G.G. (1981). "Model degradation effects on sensor failure detection," Proc. 1981 Joint Autmat. Contr. Conf., Charlottesville, Virginia.

Lou, X.C., A.S. Willsky, and G.C. Verghese (1986). "Optimally robust redundancy relations for failure detection in uncertain systems," Automatica, vol. 22, pp. 333-344.

Mehra, R.K. (1970). "On the identification of variances and adaptive Kalman filtering," IEEE Trans. Automat. Contr., vol. AC-15, pp. 175-184.

Mehra, R.K. (1972). "Approaches to adaptive filtering," IEEE Trans. Automat. Contr., vol. AC-17, pp. 693-698.

Mehra, R.K. (1976). "Optimization of measurement schedules and sensor designs for linear dynamic systems," IEEE Trans. Automat. Contr., vol. AC-21, pp. 55-64.

Mehra, R.K. and J. Peschon (1971). "An innovations approach to fault detection and diagnosis in dynamic system," Automatica, vol. 7., pp. 637-640.

Molinari, B.P. (1973). "The stabilizing solution of the algebraic Riccati equation," SIAM J. Contr., vol. 11, pp. 262-271.

Molinari, B.P. (1975). "The stabilizing solution of the discrete algebraic Riccati equation," IEEE Trans. Automat. Contr., Vol. AC-20, pp. 396-399.

Mood, A.M., F.A. Graybill, and D.C. Boes (1974). Introduction to the Theory of Statistics, McGraw-Hill, New York.

Myers, K.A. and B.D. Tapley (1976). "Adaptive sequential estimation with unknown noise statistics," IEEE Trans. Automat. Contr., vol. AC-21, pp. 520-523.

Neal, S.R. (1967). "Linear estimation in the presence of errors in assumed plant dynamics," IEEE Trans. Automat. Contr., vol. AC-12, pp. 592-594.

Neethling, C. and P. Young (1974). "Comments on "Identification of optimum filter steady-state gain for systems with unknown noise covariances"," IEEE Trans. Automat. Contr., vol. AC-19, pp. 623-625.

Nishimura, T. (1966). "On the a priori information in sequential estimation problems," IEEE Trans. Automat. Contr., vol. AC-11, pp. 197-204.

Nishimura, T. (1967a). "Correction to an extension of "On the a priori information in sequential estimation problems"," IEEE Trans. Automat. Contr., vol. AC-12, pp. 123.

Nishimura, T. (1967b). "Error bounds of continuous Kalman filters and the application to orbit determination problems," IEEE Trans. Automat. Contr., vol. AC-12, pp. 268-275.

Nishimura, T. (1970). "Modeling errors in Kalman filters," in Theory and Application of Kalman Filtering, Leondes, C.T., Ed., AGARDograph 139, pp. 89-103.

Nobel, B. and J.W. Daniel (1977). Applied Linear Algebra, Prentice-Hall, Englewood Cliffs, New Jersey.

Ohnishi, K. (1980). "Direct recursive estimation of noise statistics," in Control and Dynamic Systems, vol. 16, Leondes, C.T., Ed., Academic Press, New York., pp. 249-297.

Pachter, M. and T.E. Bullock (1977). "Ordering and stability properties of the Riccati equation," Nat. Res. Inst. for Math. Sci., Rep. WISK 264, Pretoria, South Africa.

Patel, R.V. and M. Toda (1980). "Bounds on performance of nonstationary continuous-time filters under modelling uncertainty," Automatica, vol. 20, pp. 117-120.

Pearson, E.S. and H.O. Hartley (1972). Biometrika Tables for Statisticians, vol. II, University Press, Cambridge, England.

Poubelle, M., I.R. Petersen, M.R. Gevers, and R.R. Bitmead (1986). "A miscellany of results on an equation of Count J.f. Riccati," IEEE Trans. Automat. Contr., vol. AC-31, pp. 651-654.

Price, C.F. (1968). "An analysis of the divergence problem in the Kalman filter," IEEE Trans. Automat. Contr., vol. AC-12, pp. 699-702.

Rosenbrock, H.H. (1963). "The stability of linear time-dependent control systems," J. Electron. Contr., vol. 15, pp. 73-80.

Roy, S.N. (1953). "On a heuristic method of test construction and its use in multivariate analysis," Ann. Math. Stat., vol. 24, pp. 220-238.

Roy, S.N. (1957). Some Aspect of Multivariate Analysis, John Wiley & Sons, New York, 1957.

Rudin, W. (1976). Principles of Mathematical Analysis, McGraw-Hill, New York.

Sangsuk-Iam, S. and T.E. Bullock (1987a). "Behavior of the discrete-time Kalman filter under incorrect noise covariances," Proc. 26th IEEE Conf. on Decision and Control, Los Angeles, California.

Sangsuk-Iam, S. and T.E. Bullock (1987b). "Analysis of continuous-time Kalman filtering under incorrect noise covariances," Automatica, to be published.

Schlee, F.H., C.J. Standish, and N.F. Toda (1967). "Divergence in the Kalman filter," AIAA J., vol. 5, pp. 1114-1120.

Skoog, R.A. and C.G.Y. Lau (1972). "Instability of slowly varying systems," IEEE Trans. Automat. Contr., vol. AC-17, pp. 86-92.

Tabuchi, T., T. Soeda, and T. Yoshimura (1978). "On the estimation of noise covariances in linear discrete-time systems," Int. J. Syst. Sci., vol. 9, pp. 531-542.

Tajima, K. (1978). "Estimation of steady-state Kalman filter gain," IEEE Trans. Automat. Contr., vol. AC-23, pp. 944-945.

Toda, M. and R.V. Patel (1978). "Performance bounds for continuous-time filters in the presence of modeling errors," IEEE Trans. Aerosp. Electron. Syst., vol. AES-14, pp. 912-918.

Toda, M. and R.V. Patel (1980). "Bounds on estimation error of discrete-time filter under modeling uncertainty," IEEE Trans. Automat. Contr., vol. AC-25, pp. 1115-1121.

Tse, E. and H.L. Weinert (1975). "Structure determination and parameter identification for multivariable stochastic linear systems," IEEE Trans. Automat. Contr., vol. AC-20, pp. 603-613.

Vidyasagar, M. (1978). Nonlinear Systems Analysis. Prentice-Hall, Englewood Cliffs, New Jersey.

Weiss, I.M. (1970). "A survey of discrete Kalman-Bucy filtering with unknown noise covariances," AIAA Guidance, Control and Flight Mechanics Conf., Santa Barbara, California.

Willsky, A.S. (1976). "A survey of design methods for failure detection in dynamic systems," Automatica, vol. 12, pp. 601-611.

Willsky, A.S., J.J. Deyst, and B.S. Crawford (1975). "Two self-test methods applied to an inertial system problem," J. Spacecraft, vol 12, pp. 434-437.

Wong, E. and B. Hajek (1985). Stochastic Processes in Engineering Systems, Springer-Verlag, New York.

BIOGRAPHICAL SKETCH

Suwanchai Sangsuk-Iam was born in Samutsakhon, Thailand, in 1960. He received the B.E. degree with first class honors from Chulalongkorn University, Bangkok, Thailand, in 1981 and the M.E. degree from the University of Florida, Gainesville, in 1983, all in electrical engineering.

Upon his graduation from Chulalongkorn University, he received a medal from King Pumipol Adulyadech in recognition of his four-year academic performance. In 1984, he was with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, where he was awarded the Nonresident Tuition Scholarship and the Eugene C. and Mona Fay Gee Scholarship after having studied there for a semester. Since August 1985, he has been a research assistant and a teaching assistant in the Department of Electrical Engineering, University of Florida.

His present research interests are in the areas of control, estimation, failure detection, identification, sensor fusion, signal processing, and time series.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

T.E. Bullock

T.E. Bullock, Chairman
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Giuseppe Basile

G. Basile
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

H.A. Latchman

H.A. Latchman
Assistant Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Z.R. Pop-Stojanović

Z.R. Pop-Stojanovic
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark C.K. Yang

M.C.K. Yang
Professor of Statistics

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December, 1987

Hubert A. Bevis
Dean, College of Engineering

Dean, Graduate School